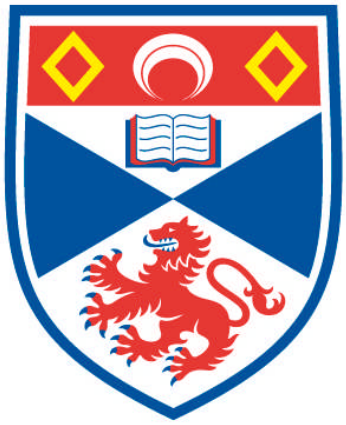
THE SOURCE OF MODAL TRUTH

Ross P. Cameron

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



2006

The Source of Modal Truth

Ross P Cameron

PhD
University of St Andrews

Submitted March 2006

Abstract

This thesis concerns the source of modal truth. I aim to answer the question: what is it in virtue of which there are truths concerning what must have been the case as a matter of necessity, or could have been the case but isn't.

I begin by looking at a dilemma put forward by Simon Blackburn which attempts to show that any realist answer to this question must fail, and I conclude that either horn of his dilemma can be resisted. I then move on to clarify the nature of the propositions whose truth I am aiming to find the source of.

I distinguish necessity de re from necessity de dicto, and argue for a counterpart theoretic treatment of necessity de re. As a result, I argue that there is no special problem concerning the source of de re modal facts. The problem is simply to account for what it is in virtue of which there are qualitative ways the world could have been, and qualitative ways it couldn't have been.

I look at two ways to answer this question: by appealing to truthmakers in the actual world, or by appealing to non-actual ontology. I develop a theory of truthmakers, but argue that it is unlikely that there are truthmakers for modal truths among the ontology of the actual. I look at the main possibilist ontology, David Lewis' modal realism, but argue that warrant for that ontology is unobtainable, and that we shouldn't admit non-actual possibilia into our ontology.

I end by sketching a quasi-conventionalist approach to modality which denies that there are modal facts, but nevertheless allows that we can speak truly when we use modal language.

I, Ross Cameron, hereby certify that this thesis, which is approximately 75 500 words in length, has been written by me, that it is a record of the work carried out by me and that it has not been submitted in any previous application for a higher degree.

[redacted] 2/3/06

I was admitted as a research student in September 2002 and as a candidate for the degree of PhD in September 2003; the higher study for which this is a record was carried out in the University of St Andrews between 2002 and 2005

[redacted] 2/3/06

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree

[redacted]

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker.

[redacted] 2/3/06

Contents

Introduction

> "The philosophical problem of necessity is twofold: what is its source, and how do we recognise it." — Michael Dummett[1]

This thesis is concerned with the first of Dummett's demands: to identify what it is in the world in virtue of which the world could have been some ways and not other ways; what are modal truths true in virtue of? Prima facie, the demand is a good one; for it is somewhat puzzling *how* a proposition can be made necessary, or *how* some falsehood can be made possible. The way things are seems, on the face of it, only to account for what is *true*, not for what *must* or *could* be true. As Ted Sider says[2] "Whether something *is* a certain way seems unproblematic, but that things might be otherwise, or must be as they are, seems to call out for explanation." The reason being that modal facts "point beyond themselves"[3] in the same way as tensed facts or dispositional facts do. What would be desirable is to locate the source of the modal in the categorical, and it will be the aim of this thesis to assess the prospects for doing so.

Now whenever one asks the question 'what is the source of the Φ-facts?' the answer 'nothing' is available. That is, it is always an option to accept some facts as primitive. It is a working hypothesis of this thesis that we can do better than this in the case of modality, and so I will be rejecting modal primitivism. Let me make clear, however, that I am not insisting on a conceptual reduction of the modal. Some modal *concept* may, for all I am presupposing, have to be taken as primitive. What I am presupposing is that whether or not modality is conceptually primitive it is not, for want of a better term, metaphysically primitive. That is, I am presupposing that if

[1] [36, p169]
[2] [137, p184]
[3] [ibid. p185]

there are modal truths then there is something in virtue of which the modal truths are true, even if the concept of modality is not to be analysed in terms of the concept of this something. (The most familiar way of providing a metaphysical grounding for truths without thereby providing a conceptual reduction of the truths in question is by providing *truthmakers* for those truths. We will look at truthmaker theory in chapter 4.)

Why reject modal metaphysical primitivism? There is something deeply suspicious about the idea that the world *just could* have been different in some ways and not in other ways, and that that is the most that can be said about the matter. The suspicion is analogous to the suspicion about the idea that the past *just was* a certain way, and that there is nothing more to be said about *why* — in virtue of what — it was that way, as opposed to some other way. That is a common suspicion against presentism — the doctrine that only present things exist. The presentist, seemingly, has to take truths about the past and future as brute truths, because they cannot ground their truth in the existence and properties of past and future objects as the eternalist can. This analogy, between the disadvantage of modal metaphysical primitivism and presentism, has been made by Ted Sider. He says[4]

> The argument against allowing the presentist to 'cheat' by invoking primitive properties like *previously containing dinosaurs*, or by invoking the tenses themselves as primitive, is that . . . the cheater is unwilling to accept an ontology robust enough to bear the weight of the truths he feels free to invoke. What seems common to all the cheats is that irreducibly *hypothetical* properties are postulated, whereas a proper ontology should invoke only *categorical*, or occurrent, properties and relations. Categorical properties involve what objects are actually like, whereas hypothetical properties 'point beyond' their instances. . . The distinction between categorical and hypothetical is admittedly elusive, though it seems to get at the core of what is wrong with the dubious ontologies. . . [The] argument . . . would work just as well against taking modal operators as primitive, for modal notions are paradigmatically hypothetical.

[4][135, p40-1]

And elsewhere[5]

> Accepting necessity or possibility as a primitive feature of reality would be like accepting tensed facts as primitive, or accepting dispositions as primitive, or accepting counterfactuals as primitive. While some are willing to make these posits, others seek to reduce 'hypothetical' notions to 'categorical' notions — notions which are in a sense 'self-contained' and do not 'point beyond themselves' as the hypothetical notions do.

I share Sider's suspicion against primitive tensed facts, and agree that it should rightly carry over to primitive modal facts. To be sure, the cases are not precisely analogous: accepting tensed facts as primitive seems to lead to a violation of the doctrine that truth supervenes upon being (since it seems that all the present objects could be exactly as they are in a world in which the past had been a different way), but accepting modal facts does not (or at least need not). Nevertheless the analogy, while not perfect, is still good; and it would be better to avoid taking modal facts as primitive if possible.

I said that it was Dummett's first demand that is my concern here, but his second demand will also play an important role for I take it to be a necessary condition for the acceptability of any account of the source of modal truth that it cohere with an acceptable epistemology of modal truth. The metaphysical story about what modal facts are true in virtue of must not make it a mystery how it is we come to know those facts. This constraint arises from a wish to respect Peacocke's Integration Challenge — the challenge to integrate the metaphysics and the epistemology in a particular area. I will say more about the Integration Challenge in chapter 5, where I will complain that Lewisian modal realism does not meet it and reject it for that reason.

Here is an outline of the chapters of this thesis.

Chapter 1: Blackburn's dilemma

Simon Blackburn raises a dilemma for any account of the source of necessity: either the facts being appealed to in order to explain the necessity of some fact are themselves necessary or they are contingent. If necessary then the question has not been answered because the source of the

[5][137, p185]

necessity of this new fact must be explained, but if contingent then this undermines the necessity of the original fact. In this chapter I discuss Blackburn's dilemma and a response to it by Bob Hale. I conclude that each horn of the dilemma can be resisted, although I postpone choosing a horn until the conclusion.

Chapter 2: Metaphysical Necessity

In this chapter I clarify the type of facts I am attempting to explain by considering two ways of thinking about the essential/accidental properties distinction and the necessary a posteriori. I contrast an ontological understanding with a deflationary understanding and argue for deflationism. According to the deflationist the only modal facts are those concerning the qualitative ways the world could have been: once we've got these there are no further questions as to the source of essentialist claims or the necessary a posteriori.

Chapter 3: Modality de re

One must distinguish between modality de dicto and modality de re, and the question as to the source of each is equally pressing.[6] In this chapter I defend Lewis' counterpart theoretic analysis of modality de re as the best deflationary account of the essential/accidental properties distinction. This chapter also contains a case study of the essentiality of origin, arguing that essentialist attempts to establish this doctrine have failed and that only the counterpart theorist can account for how we could know the doctrine to be true. I take counterpart theory to be advantageous and accept it on those grounds, and show the consequences for the source of modality de re.

Chapter 4: Truthmakers

In this chapter I examine the theory of truthmakers. This theory proposes to answer Dummett's question by finding some things in the world that makes modal claims true: a thing (or some

[6]Warning: there is a usage of 'de dicto' and 'de re' according to which what separates them is precisely their source — de dicto necessity has its source in language, while the source of de re necessity is worldly. Given the etymological roots of the expressions this is probably the more correct usage. However, it is more common now for the distinction to turn on whether we are quantifying into modal contexts or not. De re necessity, on this view, is when we attribute necessary properties to a thing (as in 'it is de re necessary of A that it is Φ'), whereas de dicto necessity is when we attribute necessity to a proposition (as in 'it could not have failed to be the case that p'). I am using 'de re' and 'de dicto' in this latter sense. So when I ask what is the source of necessity de re it is not because I do not understand the meaning of 'de re'; rather, I am asking what it is in virtue of which an object has some of its properties as a matter of necessity. That may or may not turn out to be the same as what makes it true that the proposition p couldn't be false (and either may turn out to have their source in the world or in language).

things) for each modal claim such that the modal claim is true in virtue of the existence of that thing (or those things). I look at the theory of truthmakers in general, responding to objections to it until I am satisfied that I have found the most promising version of the theory, and then I ask whether that theory has anything to tell us concerning modal truthmakers. I conclude that the prospects for finding truthmakers for modal truths in the ontology of the actual are bleak, and that if there are truthmakers for modal truths they must be non-actual entities.

Chapter 5: Lewisian Realism

Having accepted Lewis' analysis of de re modality in chapter 3, I turn my attention in this chapter to his analysis of modality de dicto in terms of truth at some world. Lewis' theory locates the source of necessity in merely possible ontology, and so modal truths no longer 'point beyond' reality in the way Sider finds worrying. Modality would simply be a matter of how things are: merely how things are at other worlds. True, modal truths still 'point beyond' *our world*; but there is nothing special about our world other than that it is the one we happen to be in, and there is no more mystery here than in the way that when we talk about what is happening "over there" we 'point beyond' where we are currently speaking.

Ultimately, though, I reject Lewisian realism on epistemological grounds; I conclude it has no good story to tell about how we can come to know that there are non-actual worlds, or about how we could come to know what went on at them if there were any, and hence that it violates the epistemological constraint. I consider John Divers' modal agnosticism which accepts the Lewisian analysis of modality and avoids the problems I present for it, but I conclude that the limits on modal knowledge the theory entails are unacceptable.

Chapter 6: Conclusion — Quasi-Conventionalism

If we have no grounds to accept non-actual ontology, and if actuality does not provide truthmakers for modal truth, then we seem in a bind: either modal truth is brute, or there is no modal truth. Since I am working on the assumption that primitivism is false I seem forced to say that there is no modal truth. But that seems to entail eliminativism: that our folk theory of modality is false.

In this final chapter I attempt to show how one might go about rejecting this last seeming entailment. I put forward an idea whereby there are no modal truths but where one can still

speak truly when we modalise. The idea is a quasi-conventionalist account of modality, in the spirit of the discussion in chapter 2.

Acknowledgements

I would like to thank the Carnegie Trust for the Universities of Scotland for financial support throughout the writing of this thesis.

I would like to thank my supervisor, Professor Crispin Wright, and my project leader and mentor, Professor Bob Hale. Whatever possible worlds there are, I doubt I could have found a better pair to work with in any of them.

I owe special thanks to my wife and fellow Arché graduate student Elizabeth Barnes, for her constant encouragement, support, and for her extremely helpful comments on every stage of this thesis.

I would also like to particularly thank my office-mate Robert Williams. My many conversations with him throughout our time at St Andrews have improved this thesis, and my philosophical ability generally, immensely.

Aside from those whom I have already thanked, thanks are also due to the following members and associates (past and present) of Arché, the AHRC centre for the philosophy of Logic, Language, Mathematics and Mind: John Divers, Frank Jackson, Carrie Jenkins, Daniel Nolan, Josh Parsons, Agustin Rayo and Sonia Roca. I would also like to thank Katherine Hawley, Andrew McGonigal and Stephen Read.

Most of section 3.3 was published as 'A Note on Kripke's Footnote 56 Argument for the Essentiality of Origin' in *Ratio* Vol.XVIII No.3 (2005) p262-275.

The ideas (but not the content) of the last couple of paragraphs of section 3.3 were the basis behind my joint publication with Sonia Roca, 'Reply to Rohbraugh and deRosset on the Necessity of Origin', *Mind* (2005).

Section 4.1, together with some of the introduction of section 4, was published as 'Truthmaker

Necessitarianism and Maximalism' in *Logique et Analyse* (2005).

Some of section 4.2 appears in my paper 'Tropes, Necessary Connections, and Non-Transferability', forthcoming in Dialectica, (although both the section and that paper contain material that the other does not).

Material from section 5.1 comprises the first half of my paper 'Much Ado About Nothing: A Study of Metaphysical Nihilism', forthcoming in Erkenntnis.

A condensed version of section 5.3 appears as 'Lewisian Realism: Methodology, Epistemology, and Circularity', forthcoming in Synthese.

Chapter 1

Blackburn's Dilemma

Blackburn raises a dilemma for anyone who attempts to identify the source of necessity. He says[1]

> Suppose an eventual answer [to the question 'What is the source of necessity?'] cites some truth F, and so takes the form: '□A because F'. . . Now, either F will claim that something *is* so, or it will claim that something *must* be so. If the latter, there is no problem about the *form* of the explanation, for one necessity can well explain another. But . . . there will be the same bad residual 'must' . . . And there is no escape from the overall problem that way. Suppose instead that F just cites that something *is* so. If whatever it is does not *have to be* so, then there is strong pressure to feel that the original necessity has not been explained or identified, so much as undermined.

Before we proceed we should note two curious things in Blackburn's presentation of the dilemma. Firstly, Blackburn is assuming that what we will appeal to as the source of necessity is the *truth* of some proposition (F). But this is certainly not forced on us, and it would be thought by many, I think, to be undesirable. What accounts for the source of necessity, one might think, is some *thing*. This would be the truthmaker theorist's thought[2]: there is some thing, or things, in virtue of which there are necessary truths. If this is one's preferred account, Blackburn's dilemma

[1] [19, 635]
[2] See chapter 4

doesn't get off the ground: it doesn't make sense to ask whether this thing, or these things, are necessary, for they are the wrong type of thing to be classified as necessary or contingent.

Suppose, for example, one believed in a thing N, the necessity maker, such that for every necessary truth p, p is necessary in virtue of the existence of N (so the truthmaker for [Necessarily, p] is N). N accounts for the source of necessity, if there is such a thing. And if Blackburn asks whether N is necessary or contingent one should reply that it is neither, because it is not a true proposition, and it is only true propositions that are necessary or contingent. One can legitimately ask, of course, whether N is a necessary or contingent *existent*[3]: that is, whether the proposition [N exists] is necessary or contingent. But the legitimacy of this question is of no help to Blackburn, for there is no regress looming if we answer that N exists necessarily. For what makes it true that N exists necessarily? N, of course! N is the necessity maker: all the necessary truths are necessary in virtue of it. There is no sense of a "bad residual 'must'." We know the answer to all questions of the form 'Why is p necessary?' — the answer is: in virtue of the existence of N. Blackburn's regress only starts if we are forced whenever we wish to account for why there are necessary truths to appeal to some necessary truth; but if we believe in the necessity maker we need make no such appeal, we only appeal to the existence of a certain thing. Nor is the original necessity in any sense undermined if N is a contingent being. For it certainly does not follow from the contingent existence of the necessity maker that the necessary truths could have failed to be necessary. It may be necessary that had N not existed there would be something else to fulfil the necessity maker role: that is, even if the thing which is actually the necessity maker only exists contingently, it may be necessary that there be *some thing or other* in virtue of which all the necessary truths are necessary. And what would make that necessary? N, of course, for it is the necessity maker.

Nevertheless, it would be interesting if Blackburn's dilemma established that we had to believe in truthmakers to account for the source of necessity. So let us grant for the sake of argument that we wish to account for the source of necessity by appealing not to the existence of some thing but to the truth of some proposition, and ask whether Blackburn's dilemma is troubling on this assumption.

[3]And one may, of course, use the words 'Is N necessary?' to ask this; but that is not what Blackburn would be asking were he to ask the question that sets up his supposed dilemma.

The second curiosity is that Blackburn says that the form of the explanation we are considering is $\Box A$ because F. But that is the explanation of the necessity of a *particular* proposition, A, whereas one might only be concerned with the explanation for why there are necessities *in general.* To be sure, one way of explaining why there are necessities in general is to explain why there is a particular necessity; since obviously if there is a necessary proposition then there are necessary propositions. But there's no reason to suppose that an answer to the question 'why are there any necessary truths?' *must* be of the form 'because A is necessary, and A is necessary because of F'. So if Blackburn's dilemma requires that the explanation be of such a form then there is perhaps a point of resistance there.

Some might think that Blackburn's question is the right one: that the demand for an explanation of the source of necessity should be understood as a demand for an explanation for why the particular necessary truths are necessary, as opposed to an explanation for why there are necessities in general. After all, sometimes a demand for explanation is only taken to be appropriate if the facts in question could have been otherwise. And it is open to think that the fact that there are necessary truths could *not* have been otherwise even if the facts that particular truths are necessary could have been. Accepting that necessarily some propositions are necessary does not commit one to accepting that some propositions are necessarily necessary. But this can't have been Blackburn's reason for focusing on the explanation for the necessity of some particular proposition rather than for there being necessities in general because, as we will see below, the contingency horn of his dilemma makes use of the characteristic axiom of S4 and assumes that the necessary truths are necessarily necessary.

Nonetheless, let us examine Blackburn's dilemma as regards attempts to identify the source of the necessity of a particular proposition; I will flag when I think it is relevant that he is not considering the view which attempts only to explain why there are necessities as opposed to why there are some *particular* necessities.

1.1 The contingency horn

First let us concentrate on the contingency horn of Blackburn's dilemma. Why is there "strong pressure" to think that if the fact we cite to explain why there are necessary truths could itself have been false then we are undermining the necessity of the proposition(s) in question? I presume Blackburn's thought is along the following lines. If F, which is to explain why there are necessities, is contingent then it might have been false. (So far so good!) So consider the world in which it is false; in that case the source of necessity is lacking. But how could the source of necessity have been lacking? Necessity is necessity; truth in *all* worlds.

As Bob Hale has pointed out[4], this reasoning seems to presuppose one of two principles: either that if the source of necessity were lacking then the necessities would no longer be necessary, or that they might not be necessary.[5] If that is right then the contingency of the source of necessity leads to the possibility that the necessary truths are not necessary. That is, the argument has one of either the following two forms.[6]

1	(1)	$\Box p$ because q	Ass
2	(2)	(α because β) $\rightarrow \neg\beta\Box\!\!\rightarrow \neg\alpha$	Ass
1,2	(3)	$\neg q\Box\!\!\rightarrow \neg\Box p$	From 1,2 $\rightarrow$E
4	(4)	$\Diamond\neg q$	Ass
1,2,4	(5)	$\Diamond\neg\Box p$	From 3,4

Or

1	(1)	$\Box p$ because q	Ass
2*	(2*)	(α because β) $\rightarrow \neg\beta\Diamond\!\!\rightarrow \neg\alpha$	Ass
1,2*	(3*)	$\neg q\Diamond\!\!\rightarrow \neg\Box p$	From 1,2* $\rightarrow$E
4	(4)	$\Diamond\neg q$	Ass
1,2*,4	(5)	$\Diamond\neg\Box p$	From 3*, 4

5 follows from 3 and 4 or 3* and 4 according to the Lewisian semantics for counterfactuals. $\neg q\Box\!\!\rightarrow \neg\Box p$ is true at a world i (according to a system of spheres \$) if and only if either (1)

[4][47, p302-303]

[5]It is obviously false that if the source of necessity were lacking, the necessary truths would not be *true*, since there are no circumstances according to which the necessary truths would not be true.

[6]As is standard, I use 'p $\Box\!\!\rightarrow q$' to symbolise 'Had p been the case, q would have been the case' and 'p $\Diamond\!\!\rightarrow q$' to symbolise 'Had p been the case, q might have been the case.

no $\neg q$-world belongs to any sphere S in $\$_i$, or (2) some sphere S in $\$_i$ does contain at least one $\neg q$-world, and the material conditional $\neg q \rightarrow \neg\Box p$ holds at every world in S. 4 tells us that there is a sphere S_1 in $\$_i$ to which there belongs a $\neg q$ world, call it W. The material conditional $\neg q \rightarrow \neg\Box p$ must be true at every world in S_1, a fortiori it must be true at W. So $\neg\Box p$ is true at W, which means that $\neg\Box p$ is possible.

$p\Diamond\!\!\rightarrow q \equiv_{df} \neg(p\Box\!\!\rightarrow \neg q)$, so $\neg q\Diamond\!\!\rightarrow \neg\Box p \equiv_{df} \neg(\neg q\Box\!\!\rightarrow \neg\neg\Box p) \equiv \neg(\neg q\Box\!\!\rightarrow \Box p)$. So from 3* we can infer that both the sufficient conditions for the counterfactual $\neg q\Box\!\!\rightarrow \Box p$ must be false. In particular then we can infer that there is no sphere S in $\$_i$ such that there is at least one $\neg q$ world and that the material conditional $\neg q \rightarrow \Box p$ is true at every world in S. But again, from 4 we can infer that there is a sphere S* in $\$_i$ such that there is a $\neg q$ world. In that case it must be that the material conditional $\neg q \rightarrow \Box p$ is not true at every world in S*; so the consequent must be false at some world in S*. I.e. there is a world in S* at which $\neg\Box p$ is true; so $\Diamond\neg\Box p$.

As I said above, there seems then to be a tacit reliance on the characteristic principle of S4, $\Box p \rightarrow \Box\Box p$, to get trouble from the conclusion at 5. For there is no problem with it being possible that p is not necessary unless we think it follows that p is not necessary, and that is just the contraposition of the S4 axiom.

Incidentally, I don't think it matters here that Blackburn's argument concerns the explanation for the necessity of a particular proposition as opposed to an explanation for there being necessities in general. For if either 2 or 2* is correct then if that which explains why there are necessities in general is lacking then there might not have been any necessary truths. Given S4 that cannot be true; S4 says there must be some necessary truths because, in particular, the actual necessary truths must be necessary. (Actually we should be more careful: it follows from S4 that *if* there are some actual necessary truths then there are necessarily some necessary truths. But the antecedent is easily discharged, since the S4 axiom itself is an actual necessary truth if it is true.)

The commitment to the S4 axiom in generating a problem means that Blackburn is forced into accepting the principle at 2 rather than the principle at 2*, because use of the principle at 2* is incompatible with the S4 axiom. 2* says that if α is true because of β then had β not been the case α might not have been the case. Now if α is itself a necessary truth then that

principle cannot be true. Might counterfactuals are existentially committing in a way that would counterfactuals are not; to assert the might counterfactual $p\Diamond\!\!\rightarrow q$ is to commit oneself to the existence of a world at which p is true and q is true. In that case any might counterfactual with a necessary falsehood as its consequent is false, since there is no world at which necessary falsehoods are true. As Bob Hale puts it[7]

> Suppose $\Box B$. Then $\Box\neg\neg B$, whence it is (vacuously) true that $A\Box\!\!\rightarrow \neg\neg B$, *and hence* [given that $p\Diamond\!\!\rightarrow \neg q \equiv_{df} \neg(p\Box\!\!\rightarrow \neg\neg q)$] *false that* . . . $A\Diamond\!\!\rightarrow \neg B$, whatever proposition A is taken to be.

Hale concludes that the might counterfactual is no good when we are concerned with explaining necessities, it is only (if at all) good when we are concerned with explaining contingencies. Hence, since we are concerned with explaining necessities, the principle is no good in this case.

But that is too quick on Hale's part. There is an ambiguity in the phrase 'explaining necessities': it may mean that the A in 'A because B' is necessary, or it may mean that it is of the form $\Box p$. In our case we can presume only that A is of the form $\Box p$, because the dilemma concerns an explanation of the form '$\Box p$ because q'. Hale's trouble with the might counterfactual, however, only arises if A itself is necessary, i.e. if it necessary that $\Box p$, and that only follows if the S4 axiom is true. What this shows is that acceptance of the might counterfactual is incompatible with the S4 axiom; it is perfectly okay for one who embraces the contingency horn to accept the might counterfactual, but then she will deny S4 and, hence, have no problem with Blackburn's conclusion, which was that the necessities are undermined because they might not have been necessary

In summary: the principle at 2*, if true at all, is only true when α is a contingent truth. In the case we are concerned with, α is a truth of the form $\Box p$, so Blackburn can only make appeal to 2* if he thinks that truths of the form $\Box p$ are contingent, i.e. if he thinks $\Diamond\neg\Box p$. But that is, of course, to deny the S4 axiom, which says that no truth of the form $\Box p$ is contingent. So Blackburn can have 2* *or* the S4 axiom, but he can't have both; so the second version of the argument is unconvincing.

[7][ibid. p303]

Blackburn needs to rely on 2 then, which says that if α is true because of β then had β not been the case α *would* not have been the case. This counterfactual does not commit us to the existence of a world at which α is true, so assertion of 2 is compatible with the S4 principle. But, of course, to assert that counterfactual *and* the claim that there is a world at which β is false commits us to the existence of a world at which α is not true. Since one who embraces the contingency horn, by definition, is one who asserts that there is a world at which the relevant β is false, that means that one who embraces the contingency horn, if they accept the principle at 2, is also committed to thinking that there is a world at which the relevant α is false. Since the relevant α is a truth of the form $\Box p$ then, as above, the theorist who embraces the contingency horn is committed to the contingency of some necessary truth, and therefore to the denial of the S4 principle. And I suppose that must be Blackburn's problem: that one cannot hold all of the S4 axiom, principle 2, and the claim that the source of necessity may have been lacking; but if the first two of these are plausible then it must be the last that goes.

But the first two aren't plausible. Here is not the place for a debate on the S4 axiom[8], but of course it is open to the contingency theorist to deny it. But really what should be first to go out of the three incompatible principles is principle 2. It is not in general true that if that which explains some true proposition p were lacking then p would be false, because in general it is the case that p *might* have been true for some other reason. Hale gives the following example.[9] Suppose I shoot at a balloon, and you shoot at it in between the time that I shoot and the time the balloon bursts. The balloon burst because I shot it. But it is not true that had I not shot, the balloon would not have burst; it might still have been burst by your shot.

Are we entitled to assert the stronger claim: that the balloon *would* still have burst because it *would* have been burst by your shot? If so then principle 2* is false, even in the case of contingencies, because the previous would counterfactual is incompatible with the might counterfactual 'If I hadn't shot, the balloon might not have burst'. But I don't think we can make that stronger claim. It's not true that had I not shot the balloon *would* still have burst, for in those circumstances you *might* have chosen not to shoot as well, or your gun *might* have misfired, or the bullet

[8]Hale takes Blackburn's reliance on the S4 axiom to show that the conventionalist shouldn't be bothered by Blackburn's dilemma, since the conventionalist will deny S4. But I think the point should be taken more generally. It is open to *anyone* to deny S4, and thereby to blunt the contingency horn.

[9][ibid. p302]

from your gun *might* have been displaced as a result of quantum indeterminacy, or etc. Such possibilities suggest the truth of principle 2* as it concerns contingencies; if (some contingent) p is true because of q, then had q been false, p might have been false; but it might also have been true. So there is nothing we can say about what the truth of contingent propositions *would* have been had the source of their truth been lacking. I conclude that the contingency horn of Blackburn's dilemma has no bite: principle 2 is not plausible, and acceptance of principle 2* goes hand in hand with rejection of the S4 axiom, which makes Blackburn's conclusion unproblematic.

1.2 The necessity horn

The problem on the necessity horn is that if that which explains why there are necessities is necessarily true then we have a "bad residual 'must'." Now here I think it is crucial to distinguish the question 'why are there any necessities?' from the question 'why are the necessary propositions necessary?' because Blackburn's worry, and any response to it, will look quite different in either case. (The questions are obviously different. The legitimacy of the questions presupposes that an informative answer can be given. But it is a perfectly coherent possibility that an informative answer can be given as to why there are any necessary propositions at all but no informative answer given as to why some particular proposition is necessary.)

Consider first the attempt to explain, for every necessary truth, why it is necessary. The problem seems to be that we land either in circularity or an infinite regress: we can explain the necessity of p_1 by appeal to a necessary truth p_2, but now we must explain the necessity of p_2; if we make appeal to the necessity of p_1 in explaining the necessity of p_2 then our explanation is circular, so we must explain the necessity of p_2 by an appeal to the necessity of some new proposition p_3. But now, of course, we must explain the necessity of p_3. We can never make appeal to the necessity of a proposition previously mentioned on pain of circularity, so we must always make appeal to the necessity of some new proposition, which leads to regress.

Now consider the attempt to explain why there are any necessary truths. If I do this by appealing to some necessary truth, p, then the worry here seems not to be that I have invited the same question again — for a demand for an explanation as to why p is necessary is not the

same as the demand for an explanation for why there are any necessary truths — but rather that I have presupposed the truth of what I was intending to explain. I have relied on the fact that there are necessary truths to explain why there are necessary truths. It is as if I have attempted to explain why there is something rather than nothing by saying 'well my parents gave birth to me, which is why I exist; hence I exist, hence there is something rather than nothing.' That is true; but obviously it is not a good explanation, for it presupposes the truth of what is to be explained when it makes appeal to the existence of my parents. So the worry here does not so much seem to be a threat of regress or circularity, but rather that a good explanation can never get off the ground. Let me deal with this latter worry first. Bob Hale has argued that it can be dispelled because we can (without embracing the contingency horn) give explanations for why there are necessary truths which do *not* presuppose the existence of necessary truths in the way that appeal to my parents presupposes that there is something rather than nothing.

Hale[10] makes a distinction between *transmissive* and *non-transmissive* models of the explanation of the necessity of some particular proposition p by appeal to some other necessary proposition. A transmissive explanation of the necessity of p is one in which the necessity of the explanans plays an explanatory role. A non-transmissive explanation of the necessity of p, by contrast, is when it is merely the *truth* of the explanans that explains the necessity of the explanandum, *even if the explanans is indeed necessary*; i.e. that while q might indeed be necessary, it is not the *necessity* of q that explains the necessity of p but merely the *truth* of q. In that case, if I give a non-transmissive explanation of why there are any necessities I do not presuppose what is to be explained: even though I have made use of a necessary truth q in the explanation, I have not presupposed that there are any necessary truths in the explanation, because the fact that q is necessary played no role in the explanation.

An example of a transmissive explanation of the necessity of a proposition p would be a proof[11] of p from premisses that are themselves necessary.[12] A proof of p from true premisses explains why p is true; but coupled with the information that the premisses of the proof are not merely true but necessary, we have an explanation of the necessity of p. The necessity of p is explained

[10][ibid]

[11]I am using the term 'proof' somewhat loosely, so that if there are undischarged assumptions at the final line it can still count as a proof.

[12][ibid. p308-9]

by the premisses, but not merely by their truth: the necessity of the premisses plays a crucial role in the explanation, for their necessity *transfers* to the conclusion. An example Hale takes to be a non-transmissive explanation of the necessity of a proposition is as follows: the explanandum is the proposition that necessarily the conjunction of two propositions A and B is true only if A is true and B is true, the explanans the proposition that conjunction just *is* that binary function of propositions whose value is true iff both its arguments are true.[13] The explanans is necessary, to be sure; but, Hale thinks, the necessity of the explanans is not doing any explanatory work in explaining the necessity of the explanandum. What explains the necessity of the proposition that a conjunction is true only if both conjuncts are is merely the *truth* of the proposition that conjunction is the function whose value is true iff both its arguments are true; no appeal to the necessity of the latter proposition is necessary to explain the necessity of the former proposition, even though the latter proposition is indeed necessary.

If Hale is right about that then it follows that we can have an explanation for why there are *some* necessities that does not presuppose that there are necessary truths. Why are there some necessities? Well, it is necessary that the conjunction of two propositions A and B is true only if A is true and B is true; hence there is at least one proposition that is necessary, and so there are necessary truths. Of course, we made appeal to some necessity here in the explanation, so let us ask now why that proposition is necessary. Because conjunction just is the function whose value is true iff both its arguments are true. And here, according to Hale, I have not made appeal to any necessity in my explanation (even though the proposition I appealed to is in fact necessary). My explanation for why there are any necessary truths can stop here; no further demand for explanation is invited for I have appealed only to the truth of some proposition, not the necessity of any proposition. As Hale puts it[14]

> [E]ven if . . . an explanation '$\Box p$ because q' of the kind suggested cannot be correct unless (one thinks that) q is itself necessary — so that the necessity of the explanans is in a sense presupposed — it does not follow that it is presupposed in a relevant way, i.e. in a way that compromises the explanation. It would do so if the explanation worked by transmitting the necessity of the explanans to the explanandum, but that

[13] [ibid. p312]
[14] [ibid. p314]

it does not do.

In short, then, we start from the *truth* of the proposition 'conjunction is the function which yields the value true iff both its arguments are true', something which does not call out for explanation[15], and use this to explain the necessity of 'the conjunction of two propositions A and B is true only if A is true and B is true', which in turn explains the truth of $\exists p \Box p$ — the claim that there are necessary truths.

Hale's non-transmission model aims to show us how we can explain why there are necessary truths without leaving a "bad residual 'must' ", but has it anything to say about the demand to explain the necessity of the *particular* necessary truths? It seems not.[16] If my goal is to give an explanation for the necessity of each necessary truth, and I begin by explaining the necessity of some necessary proposition p by an appeal to the necessary truth q, then the demand for an explanation for the necessity of q arises *whether or not the necessity of q was relevant to the explanation of the necessity of p.* When our goal is to give an explanation of the necessity of p, for all necessary propositions p, it seems we have *no choice* but to make appeal to some contingent proposition at some stage as the explanation for the necessity of some proposition if we want to avoid circularity or regress. But how bad a conclusion is this? Suppose one wants to avoid appeal to a contingency, how worried should we be about the looming circularity or regress that Blackburn threatens us with? I think at least in some cases the answer is 'not very'. Consider, for example, the Lewisian answer to the source of necessity (which will be discussed more fully in chapter 5). Lewis says that $\Box p$ is true because p is true at every world. Now that latter fact is, according to the Lewisian, necessary[17]; but the necessity of the proposition is playing no role in the explanation of the necessity of p, merely the truth of the proposition, and so we have a non-transmissive explanation of the necessity of p, and hence of the claim that there are some necessary truths. But of course while we might at this point be able to claim that we have explained why there are necessary truths we cannot claim that we have explained the necessity of every necessary truth. In particular there is a necessary truth — at every world it is true that p — whose necessity has not been explained. What explains the necessity of this

[15] At least it does not call out for explanation in *this* context: a context in which we are aiming to explain *necessity* as opposed to *truth*.

[16] Nor does Hale claim otherwise.

[17] See Divers [31] for a Lewisian treatment of the modal status of claims concerning the plurality of worlds.

proposition? The fact that at every world it is true that at every world it is true that p. This fact is also necessary. Why? Because at every world it is true that at every world it is true that at every world it is true that p. We are clearly heading down a road with no end, but should we be worried about this regress? Is it vicious? I think not. There is no mystery about why any necessary truth is necessary. We are told the general reason why necessary truths are necessary: because they are true at every world. If modal realism is correct then no necessary truth is left mysteriously unexplained, even if we face regress when we try to explain the necessity of every necessary truth. I conclude then that there is not in general a problem with the kind of regress Blackburn threatens when we attempt to explain the necessity of every necessary truth.

I hope I have done enough now to show that Blackburn's dilemma should not be taken as showing that it is hopeless to try and identify the source of modal truth; let us now turn to examine various attempts at providing such an identification.

Let me start by saying something about Hale's suggestion above that it's necessary that the conjunction of two propositions is true only if both conjuncts are true because conjunction just *is* that binary function of propositions whose value is true iff both its arguments are true. Another example Hale gives is that it is necessary that vixens are female foxes because being a vixen just *is*, or consists in, being a female fox.[18]

I am doubtful that this is in fact an explanation of the necessity of the propositions in question. Consider the impossible world[19], w, where both A and B are true but in which the conjunction A∧B is false. Does the fact that conjunction just *is* that binary function of propositions whose value is true iff both its arguments are true explain why w is impossible? I can't see why it would. Sure, conjunction is that binary function of propositions which is true iff both its arguments are true; but in w such a function can have true arguments but yield a false value. That is impossible, of course; a function which is true iff both its arguments are true cannot possibly have true arguments but yield a false value. But is this impossibility explained by

[18] Hale op cit. p312.

[19] I think that anyone other than the Lewisian realist ought to believe that there are impossible worlds in the same sense as there are possible worlds. Impossible worlds have as much utility as possible worlds: for a start, given impossible worlds we are not forced into the conclusion that counterpossibles (counterfactuals with impossible antecedents) are vacuously true. Intuitively that is the correct result: for while 'Had $\sqrt{2}$ been rational it would have been possible to write it as m/n' appears to be true 'Had $\sqrt{2}$ been rational it would have been greater than 5' does not. On non-vacuously true counterpossibles see Nolan [96] and Vander Laan [61].

conjunction being the function it is? I think not; for it is that function in w as much as it is in the actual world. It is simply that in w it is behaving in an impossible manner. But what makes this situation an impossibility is so far left unexplained.

To put what I think is the same point slightly differently: isn't the reason w is impossible not to do with how conjunction in fact behaves, but rather with the fact that conjunction behaves the way it does *necessarily*: that as a matter of necessity a conjunction yields truth when both its arguments are true? But that is not to explain, or to give grounding to, the necessity in question: it is merely to restate it.

But more importantly, even if Hale's account adequately explains why it is necessary that vixens are female foxes or that A∧B is true if both A and B are true, his story seems unlikely to generalise to give an account of the source of necessity in general.[20] There are many propositions one might think are necessary which don't appear amenable to Hale's account. Try completing the following claims with anything that sounds both plausible and informative:

- It is necessary that nothing is both red and green all over because being red just *is* . . .
- It is necessary that God exists because God just *is* . . .
- It is necessary that all the actual abstract objects exist because abstracta just *are* . . .
- It is necessary that whenever there is a collection of simples there is an object that has exactly those simples as parts because being a collection of simples just *is* . . .
- It is necessary that if there is a universal there is something to instantiate that universal because universals just *are* . . .
- It is necessary that there are some concrete objects because concrete objects just *are* . . .
- It is necessary that nothing is wholly located at a time t and wholly located at a time t* (such that t≠t*) because persistence just *is* . . .

[20] I should point out that Hale does not claim otherwise; that is, nowhere does Hale lay claim to be offering a complete account of the source of necessity, merely to be offering an account of the source of some necessary truths. But I am looking for a complete account of the source of necessity, which is (one reason) why I am not happy to rest with Hale's proposal.

No explanations of the type Hale favours seem forthcoming here, and there are many more examples. Of course, we could say things like 'It is necessary that God exists because God just *is* that being who exists no matter what' or 'It is necessary that whenever there is a collection of simples there is an object that has exactly those simples as parts because mereological sums just are those things whose existence conditions are simply that the relevant simples exist'. But such claims seem merely to restate the necessity of the truths in question, not to ground their truth. Let me make it clear that I am not accepting that all the propositions listed above *are* necessary; but they are examples of what people have thought to be necessary truths, and I would rather have an account of the source of necessity that was compatible with the necessity of at least some of the propositions in question. Since Hale's account does not seem to generalise to them, it is worth pursuing alternatives. But before we look at alternative accounts of the source of modal truth, I want to make it clear just what type of facts I am aiming to account for. This is the aim of the next chapter.

Chapter 2

Metaphysical Necessity

What I am concerned with in this thesis is the source of *metaphysical* necessity, by which I mean to include not only truths of logic or of maths, or in general those which are analytic or a priori, but also those a posteriori necessities and truths concerning essential properties brought back into philosophical attention by, among others, Kripke.[1] But there are two ways to take these 'metaphysical' necessities, and how one is inclined to treat them may influence one's approach to the source of modal truth. One can treat these metaphysical necessities with ontological seriousness or in a more deflationary manner. Let me make clear the distinction by considering how it applies to the essential/accidental properties distinction and the necessary a posteriori in turn.

[1] [60]

2.1 Essential versus Accidental properties

Let us consider essentialism first: the theory that an object's properties divide into those it has essentially — i.e. those which it must have if it is to exist[2,3], and those which it has accidentally — i.e. those which it could have lacked.

Now suppose we accept such a distinction between an object's properties. We are then faced with the question: is the source of the distinction *ontological* or is it rather a result of how we are conceiving of the object? Is it the world that carves the distinction between essential and accidental properties or is it, in some sense, us? The distinction is perhaps made clearer by considering examples of either position. Consider the difference between Lewisian modal realism with counterpart theory and Lewisian realism with trans-world identity, where the trans-world identity in the latter is whole location in each world: the modal analogue of endurantism.[4] The counterpart theorist falls into the deflationary camp when it comes to modal properties while the modal endurantist (for want of a better term) falls into the ontological camp. For the modal endurantist an object has a property essentially iff it has it in all the worlds in which it exists; and of course that is something which is completely settled by reality, for what worlds the thing exists in is nothing to do with us, and what properties it has in each world in which it exists is nothing to do with us. For the counterpart theorist, on the other hand, an object has a property essentially iff all of its counterparts have that property. Now of course reality settles what properties the counterparts have, that has nothing to do with us; but reality does not settle just what individuals *are* the counterparts — that is a result (partly) of how we conceive of the object in question. So the counterpart theorist does not see an ontological difference between the essential and accidental properties of a thing as the modal endurantist does. For the counterpart theorist the division of properties into those had essentially and those had accidentally is a result of whichever standards of similarity are contextually salient. When you think of Jakob as the

[2]This is sometimes taken to be equivalent to 'those which it must have if it is to be the very thing it is'. But these are not equivalent unless we rule out the possibility that a thing could have been other than the thing it in fact is, which I am not prepared to rule out. (Even ruling that possibility out they might not be equivalent; Fine has argued that there are some properties which are had necessarily by things, but which are no part of what it is to be those things: for example the property of being such that 2+2=4 as had by any non-number, or the property of being distinct from the Eiffel tower, as had by anything that is not the Eiffel tower. See Fine [38].)

[3]Let us also restrict our attention to non-trivial essential properties, i.e. those properties which an object must have but which it is possible that some thing lack, so as to exclude trivial essential properties such as the property of being self-identical.

[4]For the development of such a theory see [87].

child of Bob and Sara perhaps you create a context according to which Jakob is essentially born of Bob and Sara, but perhaps accidentally male. But it is not that Jakob is more deeply tied to the property of being born of Bob and Sara than the property of being male; it is simply that in this context who Jakob's parents are is a more salient feature than his sex, and as a result an individual who differs in origin is too dissimilar from Jakob to be his counterpart while a difference in sex is acceptable. Had we instead thought of Jakob as the *son* of Bob and Sara we would perhaps have created a context in which Jakob was essentially male and essentially born of Bob and Sara, whereas if we had thought of him simply as 'that person over there' perhaps we would have created a context in which he is both accidentally male and accidentally born of Bob and Sara.

The difference between the deflationary and ontological accounts here is a result of how the two theorists think of representation de re. For the counterpart theorist a world represents Jakob as existing by containing an individual distinct from Jakob who is similar to him in relevant ways; and the reason there is no ontological difference between the essential and accidental properties is that what counts as a relevant respect of similarity is not a matter that is settled wholly by how things are but that is settled in part by how we think of Jakob. Whereas if Jakob is represented as existing at a world by being wholly present there, as the modal endurantist maintains, then it is true or false independently of how we think of him whether or not he is born of Bob and Sara at every world.

Consider another example of the distinction. Judith Jarvis Thomson, when considering the relationship between the statue and the lump of clay that constitutes it, accounts for one intuitive difference between the modal properties of the statue and the clay — namely that the statue could lose some parts and still remain that very statue but that the clay could not lose any parts and remain the very same lump of clay — by claiming that the clay and the statue bear a different ontological relationship with their parts. She says "In short [the clay] is more tightly tied to its parts than [the statue], and *that is the ontological difference* between them."[5]

Laurie Paul considers a view that she notes is similar to Thomson's. The view Paul considers is a bundle theory whereby objects are mereological sums of tropes, but there is more than one

[5][144, p157], emphasis added

fusion relation. If an object O $fuses_1$ the property P then it has P accidentally whereas if it $fuses_2$ P then it has it essentially. So if O_1 is essentially F and accidentally G and O_2 essentially G and accidentally F the difference is due to the different ways in which F and G are a part of O_1 and O_2.[6] This is a view according to which there is an ontological difference between the accidental and essential properties of a thing; the distinction is a result of the different ways in which an object can fuse a property, and that is something that reality decides, not us. Contrast that with a view that says that there is only one way for an object to fuse a property — only one fusion relation — but that some fusions of properties are more psychologically salient to us than others, and it is these ones that are the essential properties. Here there is no ontological difference between the essential and accidental properties; the distinction is merely a projection on our parts.

I am sympathetic to a deflationary account of the essential/accidental properties distinction, but I will leave a defence of deflationism in this respect until chapter 3 when I will defend what I think is the best way of being a deflationist: counterpart theory. For now let us turn to look at the distinction as it applies to the necessary a posteriori.

2.2 The Necessary A Posteriori

I have contrasted ontological with deflationary accounts of the distinction between essential and accidental properties. A similar distinction arises with respect to the a posteriori necessities, e.g. water is H2O, gold has atomic number 79 etc. According to the deflationary view, the necessary a posteriori is primarily a linguistic phenomenon rather than a metaphysical one. The phenomenon of the necessary a posteriori arises, on this view, from considerations about how we properly *describe* certain situations. The lesson we should take from Kripke and Putnam, so the view goes, is that we can only know a posteriori that certain sentences truly describe every possible situation. But that is not because we need empirical evidence to discover that the state of affairs that those sentences claim to obtain obtains necessarily; rather it is because we need empirical evidence to discover just what state of affairs those sentences claim to obtain. If I utter

[6] [100, p185]

the sentence 'there is some non-H2O water around' then I utter a necessary falsehood. Why? Because the state of affairs that sentence claims to obtain is one where there is some of the H2O substance in its liquid form that is not H2O. That state of affairs cannot obtain, and I know a priori that it cannot obtain; but what I cannot know a priori is that that is the state of affairs said to obtain by that sentence. To know that I need to discover the chemical composition of water: I need to know that water is the H2O substance in its liquid form. There is no claim here that 'water' *means* 'the H2O substance' — that is not plausible, since we understood the term 'water' before we knew anything of chemical theory; rather the claim is simply that knowing the meaning of the words in a sentence is not always sufficient to know what state of affairs is said to obtain by that sentence. Sometimes we must also know some a posteriori facts about the things that our words apply to.

The phenomenon of the necessary a posteriori arises, according to the deflationist, because there are certain *sentences* which truly describe every possible world but which can only be discovered to do so by empirical means. The phenomenon of the necessary a posteriori does *not* require us to admit that there are *worlds* which are impossible which we can only discover to be impossible by empirical means. Our opinion on what worlds are (metaphysically) possible did not change post Kripke/Putnam, says the deflationist; all that changed was our opinion on how some of those worlds are correctly *described*, which has an impact on our opinion as to what *sentences* are necessary (or, if you prefer, what sentences express necessary truths).

I have a great deal of sympathy for the deflationist story here. The most convincing consideration, to my mind, that water couldn't have been anything other than H2O is the Twin Earth thought experiment. We were asked to consider a world in which there is a counterpart of Earth where the clear, odourless, colourless liquid that falls from the sky, makes plants grow, quenches thirst etc is not H2O but XYZ, and then asked whether or not this is a world which contains water. We answer that it isn't. Now what is crucial is that what the Twin Earth example elicits from us is an opinion about how we should *describe* the world in question. It's not, or so it seems to me, that one of the ways we thought a possible way the world could be turned out not to be possible; it's that it turned out that we were misdescribing one of the possible ways for the world to be as a way according to which water was not H2O, when in fact the proper description is

that it is a way for the world to be in which there is no water, but merely twater. Our conception of the underlying modal facts, then, is no different post Kripke/Putnam from how it was pre Kripke/Putnam; all that changed was our opinions on how to properly describe those facts: on what sentences express truths as a result of the modal facts being what they are. What Kripke and Putnam did is convince us that some of the possibilities were being misdescribed; and the phenomenon of the necessary a posteriori arises because to know whether or not a possible situation is properly described as one where there is water we need to do some empirical work — we need to discover the *actual* chemical composition of water.

Contrast this with a view which places ontological importance on the necessary a posteriori. Consider, for the sake of illustration, a view according to which there are two distinct universals: a simple universal of being water and a structured universal of being H2O. Something is water iff it instantiates the being water universal, and it is H2O iff it instantiates the structured universal of being H2O. Now certainly there is no *logical* contradiction in something instantiating one and not the other but, according to this view, such a situation is metaphysically impossible, perhaps because of the natures/essences of the universals in question. On this view the necessity of water being H2O is not simply a result of how we describe the space of possible worlds — whether or not something instantiates the water universal is independent of whether or not we would call that something 'water'. This view claims that there are two distinct universals which as a matter of necessity must be either both instantiated by a substance or both uninstantiated. In contrast to the deflationary theorist this theorist thinks that we did learn something about what worlds were possible from Kripke/Putnam, not just about how to correctly describe the space of possible worlds. According to this theorist we can only know a posteriori what situations are possible. A priori reasoning cannot reveal that the world where something instantiates the water universal but not the H2O structured universal is an impossible world.

That way of taking the necessary a posteriori with ontological seriousness is not very attractive, but at the moment I am merely aiming to illustrate the distinction, and I hope I have done that. That particular theory is not very attractive because it is not very plausible to think that there are two distinct universals, the water universal and the H2O universal. If you believe in universals it is likely that you only believe in simple universals corresponding to sparse properties and

complex universals constituted from them; you will probably not believe in a simple universal of being water. If you want to take abundant properties such as being water with ontological seriousness, chances are you take them to be something like functions from possible worlds to sets of individuals. Assume that is the case. So the property of being water is a function such that when it takes a world w as its argument it yields as its value the singleton of the substance water, or perhaps the set of all the water molecules, in w. So there are indeed two distinct properties — the structured universal being H2O is not identical to this function — but nevertheless this is a road back to deflationism. For the reason that nothing can be water and not H2O according to this view is not plausibly to be taken as the result of some necessary connection between the structured universal being H2O and the function being water, as is the case according to the previous view, but rather because our usage of the term 'water' is such that the term would not apply to any function which took you from a world to a thing in that world that did not instantiate the H2O structured universal. And so the peculiarity of the necessary a posteriori remains wordy rather than worldly.

That our usage of the term 'water' imposes such a constraint on what function is the property of being water is only knowable a posteriori, for it is obviously not analytic of the term 'water' that a function is the property of being water only if it never takes you from a world to a non-H2O substance in that world; rather the view would be that it is analytic of the term 'water' that a function is only the property of being water if it never takes you from a world to a substance in that world which has a different chemical composition from the actual watery stuff. Since it is an empirical matter what that chemical composition is, it is an empirical matter what the constraint is that is placed on whether or not a function deserves to be called the 'being water' property. Again then, according to this view we know a priori the space of possibilities but we do not know a priori how to describe the space of possibilities, because in order to know that the function that takes you from the Twin Earth world to the watery XYZ stuff is not the property of being water you need to know that water is not in fact XYZ.

2.2.1 Kripke versus Putnam

The ontological/deflationist distinction as it concerns the necessary a posteriori has its representatives in Kripke and Putnam respectively. For Kripke, it was a metaphysical discovery that the essence of water is that it is H2O; for Putnam the necessity of water being H2O said more about how we use the term 'water' than about any metaphysical nature of the substance itself. Putnam wanted to accept the necessary a posteriori but to "assimilate [Kripke's] *metaphysical* intuitions to . . . *linguistic* intuitions."[7] To arrive at "a theory which was related to Kripke's, but which was stripped of metaphysical assumptions to the point where *Carnap* might have accepted it."[8]

What is the source of the difference between Kripke and Putnam? In this case it is down to a difference in how they thought about trans-world identity.[9] For Kripke, remember, no criteria of trans-world identity are ever needed. It is not as if worlds merely give a qualitative description of how things are and we are then faced with the task of identifying what, if anything, in that world is water; rather, we simply *stipulate* that such-and-such stuff in this world is water. So we can stipulate that the content of the counterfactual situation under question is one in which there is water that is not H2O. What we cannot do, of course, is stipulate that such a situation is possible; whether or not the situation in question is one which could obtain or not is a result of the *essence* of water. So for Kripke questions about trans-world identity do not arise; we can simply stipulate the content of counterfactual situations, and then their possibility or impossibility is determined by the natures of the things we have stipulated to exist in those situations. That is a thoroughly ontological understanding of the necessary a posteriori then; the necessity of 'water is H2O' is a result of the nature of water, a nature it has independently of how we think about water.

Putnam will have none of this. For Putnam there are no 'natures' or 'essences' that determine the extent of what is possible for certain things or substances. Rather, questions of trans-world identity are settled partly by convention; it is in some sense up to us whether or not the non-H2O stuff is to count as water. He says[10]

---

[7] [113, p64]
[8] ibid. p63-64
[9] I am very indebted here to Bob Hale's article on Putnam's views concerning metaphysical necessity, [50].
[10] Putnam op cit. p67

> [For Kripke] statements about "this table" [or about water, gold, this person etc] require . . . an intuitive knowledge of what is "essential" to the table — an intuitive grasp of the limits of the possibilities in which the hypothetical object would bear the primitive logical relation "=" to the table I am pointing to [or to the substance water, or etc]. [In contrast] Criteria of table-identity are conceived of (by me, anyway) as to some extent *up to us*. [Whereas] Facts about "=" are not (in Kripke's view, anyway) at all up to us.

Putnam says the necessity of water being H2O is quasi-conventional. It is conventional, as I understand him, insofar as it is a matter of convention whether Twin Earth is taken to be a water world or not; it is merely *quasi*-conventional insofar as the modal *facts* do not hold true in virtue of our conventions, merely the correct description of those facts.[11] This is in direct opposition to Kripke, who thinks that the essence of water determines that the XYZ stuff in the Twin Earth world is not water: it is reality that settles that question, and not convention. Kripke agrees with Putnam, of course, that we could have used the term 'water' in such a way that XYZ would have deserved the label 'water' provided it has the surface qualities of water. But Kripke, unlike Putnam, thinks that we would thereby be neglecting a fact about the essence of the natural kind that we in fact use the term 'water' to pick out. For Putnam, and for me, there is no such fact about the essence of water.

Putnam's idea is close to the counterpart theorists' idea in many respects. According to the counterpart theorist there is no mind independent fact whether or not a non-actual possible world w represents an actual individual O as existing there. w represents some individual existing — that much has nothing to do with us; but whether or not the individual is a representation of O is up to us, because this individual is only a representation of O if it is similar in relevant respects to O, and what respects are relevant is up to us, it is not settled by the mind-independent reality. Similarly, w may represent there being some water like substance, and it may represent this water like substance as having the chemical constitution XYZ. But there is no mind independent fact

[11] Of course both Kripke and Putnam — both the ontologist and the deflationist — agree that there is some sense in which the correct description of the modal facts is a matter of convention, for conventions determine that our words mean what they do. But the issue is over whether this is the entirety of the role convention plays. The ontologist thinks that once the meaning facts are settled reality thereby settles whether the sentence 'water is H2O' expresses a necessary truth. The deflationist thinks that convention plays a further role: in determining what substances at other worlds count as water.

of the matter whether w thereby represents *water* as being XYZ. Whether or not this substance is a representation of water depends on what features of water must be had by a substance in order for it to count as representing water, and what features those are depends on the conventions governing our usage of the term 'water' (or perhaps on the conventions governing the term 'substance'). And, the story goes, the conventions that are in fact in place governing the term 'water' are such that similarity in chemical composition is a necessary condition for some substance at a world w to represent water.

As I said, my sympathy lies with Putnam — with the deflationist — and against Kripke — against the ontologist — on this. But it is important to recognise the limited sense in which this position is deflationary. This is not an eliminativist position; there has been no deflation of modal facts. It is being accepted (or at least, nothing has been suggested to the contrary) that there are mind-independent objective modal facts. The issue is solely over whether those facts in and of themselves settle the question of whether water could have been XYZ, gold the element with atomic number 76, Aristotle a dog, and so on. Kripke says they do, Putnam says they don't; and I agree with Putnam. The modal facts determine only the *qualitative* ways things could have been; it is a further question, and one that is settled as a matter of convention rather than by some mind-independent objective facts, as to what sentences come out necessary (or express necessary truths) as a result of the modal facts being as they are.

My view, then, is that if there is a space of mind-independent modal facts then they are facts that represent, contra Kripke, the *qualitative* ways things could have been. Those are the modal facts, assuming there are such facts, and those are the facts whose source I aim to identify. But those facts leave it open what sentences come out as necessary or contingent, for they do not settle whether or not in a particular world the watery stuff is actually water, or the gold-like stuff actually gold, or a particular human Aristotle. *We* settle such questions, as a matter of the conventions governing usage of the terms 'water', 'gold' and 'Aristotle'.

Chapter 3

Modality de re

In the last chapter I contrasted two ways of thinking about the distinction between essential and accidental properties: one could think of the distinction as being drawn by the world or by us; that is, one could think that there is a genuine ontological difference between having a property essentially and having it accidentally, or one could think that this distinction is drawn somehow by how we think of the object in question.

As I said then, I favour a deflationary account whereby there is no genuine ontic distinction between essential and accidental properties, and in this chapter I argue for this and for counterpart theory, which is I think the best theory to make sense of the deflationary story. Counterpart theory is a theory of representation de re. According to the counterpart theorist, a non-actual world w does not represent an actual individual A as being a certain way — being such that Φ say — at w by A existing at w and being Φ there; rather a *counterpart* of A — something *like* A in relevant respects — exists at w and is Φ there. Since what those relevant respects are is not settled by mind-independent reality but by how we think of the object, it follows that whether or not w represents A as being a certain way is partly down to us.

3.1 The Inconstancy of Modality de re

The main argument I have in favour of the deflationary account of the essential/accidental properties distinction — aside from a brute aversion to the thought that the world carves properties so — is the inconstancy of modality de re; and my main argument for counterpart theory that it tells the best story about why representation de re is inconstant. The argument basically runs as follows:

1. What is possible of an object can vary from context to context.[1]
2. What an object is identical to cannot vary from context to context.
3. From 2, if there was trans-world identity then what is possible of an object would not vary from context to context.
4. From 1 and 3, there is no trans-world identity.
5. From 4, the relation between an object O in w and an object O* in w* (such that O* represents a possible way for O to be) is a counterpart relation. (Hence, counterpart theory is true.)
6. Whether or not an object O* is a counterpart of O is not a mind-independent objective matter.
7. From 5 and 6, what essential properties an object has is not a mind-independent objective matter. (Hence, deflationism is true.)

The move from 2 to 3 is not watertight: one *could* hold the inconstancy of essentialist claims and be a trans-world identity theorist, and we shall look later at an attempt to do just that by Laurie Paul. But I think that the *best explanation* for why essentialist claims vary from context to context is that there is no identity across worlds but rather a relation of counterparthood which is context sensitive. Thus the inconstancy of essentialist claims, while not proving counterpart theory (and, as a result, deflationism) is good evidence for counterpart theory. Let me now move to argue for such inconstancy.

[1] I would like to be able to offer a theory of what contexts are, but I cannot. I will use 'context' as a technical term to pick out whatever it is that accounts for different sentences of the same type having different truth-values.

3.1.1 Multiple counterpart relations and Leibniz's Law

According to counterpart theory, an object could be Φ iff it has a counterpart at some world who is Φ (and so is essentially Φ iff all its counterparts are Φ). The most important development in Lewis' thoughts on counterpart theory was the admission of more than one counterpart relation. If there is only one counterpart relation then there is a fixed answer as to what properties an object has necessarily, or what properties it possibly has; but if there is more than one counterpart relation then whether an object is possibly or necessarily Φ only makes sense relative to some counterpart relation. So an object O is possibly/necessarily Φ *relative* to some counterpart relation C iff under the C counterpart relation some/all of O's counterparts are Φ. What counterpart relation is selected is determined by contextual factors, so this yields the immediate consequence that what properties an object possibly or necessarily has can vary from context to context — the answer is not fixed once and for all. Most importantly, context includes how we refer to an object, so different counterpart relations can be selected when we speak of the same object in different ways. Importantly, then, this means that Leibniz's law can fail for modal properties. For suppose it is true that a is necessarily Φ but that b might not have been Φ; does it follow that a and b are distinct? Only if it is one and the same context in which a is necessarily Φ and b not necessarily Φ. If it is true that a is necessarily Φ and that a=b then, provided that there is no change in context, it is true that b is necessarily Φ. In practice, however, referring to the one thing as b can often be sufficient to change the context from when we had referred to it as a; and so it will often not follow from the truth of 'a is necessarily Φ' and the identity of a and b that 'b is necessarily Φ' is true. And as a result it will not follow from the fact that a and b differ in modal properties that a and b are distinct.

We will come back to Lewis' reasons for admitting multiple counterpart relations, but for now let us look at the historical precedent for the inconstancy of modality de re.

3.1.2 Quine and the inconstancy of representation de re

The inconstancy of representation de re really goes back to Quine, although Quine used it as an argument against de re modalising: what he called the third grade of modal involvement. Let

us look at two of Quine's examples. First, consider Quine's example of the number of planets.[2] Quine senses trouble in the supposed truth of the claim that 9 is necessarily greater than 7 and the supposed truth of the claim that the number of planets could have been less than 7. For the number of planets = 9, and so by the substitutivity of identicals we infer from the first claim the negation of the second, that the number of planets is necessarily greater than 7. Consider also Quine's example of the mathematician who is also a cyclist. Qua mathematician he is, thinks Quine, essentially rational but not essentially two-legged; whereas qua cyclist he is essentially two-legged but not essentially rational. He says[3]

> Perhaps I can evoke the appropriate sense of bewilderment [concerning essentialism] as follows. Mathematicians may conceivably be said to be necessarily rational and not necessarily two-legged; and cyclists necessarily two-legged and not necessarily rational. But what of an individual who counts among his eccentricities both mathematics and cycling? Is this concrete individual necessarily rational and contingently two-legged, or vice-versa? . . . [T]here is no semblance of sense in rating some of his attributes as necessary and some as contingent.

However, with counterpart theory on the table, it seems that Quine's conclusion that de re modalising should be abandoned is too strong; isn't the most that is warranted by the above example simply that, as counterpart theory allows, what is de re possible of an object can vary depending on how one refers to that object?[4] 'X is necessarily two-legged' is true iff all the counterparts of X are two-legged; but who the counterparts of X are can vary depending on whether we refer to X as 'the cyclist' or 'the mathematician', so it is perfectly understandable, and acceptable, that 'the cyclist is necessarily two-legged' is true but 'the mathematician is necessarily two-legged' is false, even when 'the cyclist' and 'the mathematician' pick out the same person. Likewise, mutatis mutandis, for 'X is necessarily rational'. Similarly, 'X is necessarily greater than 7' is true iff all the counterparts of X are greater than 7; but '9 is necessarily greater than 7' can be true and 'the number of planets is necessarily greater than 7' false because what things count as a counterpart of that object which is both 9 and the number of the planets

[2] [121, p163],[117, p143-144]
[3] [118, p199]
[4] This Lewisian response to Quine was brought home to me by John Divers. See his [34].

can vary depending on whether we refer to that object by the expression '9' or 'the number of planets'.

The counterpart theoretic answer is of course by no means mandated; essentialist solutions[5] are available to both the Quinean examples. In the case of the number of planets example[6], we must be careful to distinguish two claims; that it is necessary that the number which numbers the planets is greater than 7, and that the number which numbers the planets is necessarily greater than 7; i.e. between

$$\Box\exists x(Nx \wedge x > 7)$$

and

$$\exists x(Nx \wedge \Box x > 7)$$

and of course there is a third option as well

$$\exists x\Box(Nx \wedge x > 7)$$

which says that there is a number which necessarily numbers the planets and is necessarily greater than 7.

It is only the second of these that the essentialist is going to accept: that the number which in fact numbers the planets, 9, is necessarily greater than 7. The first will be rejected because there could have been only 6 planets, and the last rejected because the number of planets is contingent and so there is no number which necessarily numbers the planets. But it is also only the second, neither the first nor the third, that is obtainable from a true statement via Leibniz's law. For consider the statement 'nine is necessarily greater than 7'; that is a statement that says *of* a particular number that it is necessarily greater than 7. Now substitute 'the number of planets' for 'nine' and we get 'the number of planets is necessarily greater than 7'. If that is still a de re claim then it is true; it says *of* a particular number which is necessarily greater than 7 that it is necessarily greater than 7. The statement *sounds* odd, of course; but that is presumably because we are inclined to hear it as the false de dicto claim that necessarily the number of planets is

[5] By which I mean solutions which uphold the constancy of essentialist claims.
[6] The broad line of solution discussed here was put forward by Arthur Smullyan [139].

greater than 7; the statement sounds a lot better when it is legitimately paraphrased as 'the number which in fact numbers the planets is necessarily greater than 7'. The false de dicto claim does not follow from the claim that nine is necessarily greater than 7; it *does* follow from the first and third claims above, but this is no problem because the essentialist wouldn't accept them in the first place.

Perhaps, however, Quine did not choose his example well enough. Instead of Quine's example consider a slightly different one. The set {Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus, Neptune, Pluto} necessarily contains 9 members. The set {Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus, Neptune, Pluto} is the set of planets. Therefore, by Leibniz's law, the set of the planets necessarily has 9 members. I think our pre-theoretic intuitions are far less likely to judge the conclusion here as true than they are to judge as true the de re reading of 'the number of the planets is necessarily greater than 7'. The point here is that sets may often be specified intensionally or extensionally. It seems not unnatural to me to suppose that the modal properties of the set vary depending on whether one thinks of its members via its intension or its extension. If we think of a set via its extension, as being the set with members $x_1, x_2, \ldots x_n$, then of course it seems that the set could not have members other than $x_1, x_2, \ldots x_n$; if on the other hand we think of a set via its intension, as being the set *of* the Ys, then it seems natural to think that the members of the set would vary as the Ys varied, so in a world in which the Ys were not $x_1, x_2, \ldots x_n$ but rather $x_{1*}, x_{2*}, \ldots x_{n*}$ the very set which in fact has as its members $x_1, x_2, \ldots x_n$ would have as its members $x_{1*}, x_{2*}, \ldots x_{n*}$. But it is as clear as anything that there is only one set here; I describe one and the same set whether or not I describe it intensionally or extensionally; and so if there is a difference in the modal properties of the set depending on how I describe it that must be because of the inconstancy of modality de re. That seems natural to me, but I will not try to push the point as I don't imagine it convincing many essentialists who, having had their intuitions corrupted by many years of thinking as a trans-world identity theorist, will likely deny that when we think of a set as the set of the Ys that its members will vary as the Ys vary.

As for the mathematical cyclist, the essentialist will simply deny the data Quine is working from. This person, they will hold, is not essentially rational nor essentially two-legged, no matter

how he is thought of. Plantinga, for example, points out that the following 3 statements

1): Cyclists are necessarily bipedal but not necessarily rational

2): Mathematicians are necessarily rational but not necessarily bipedal

3): Someone is both a cyclist and a mathematician

only form an inconsistent triad if (1) and (2) are read de re as

1_{DR}): Every cyclist has the property of being necessarily bipedal but lacks the property of being necessarily rational, and

2_{DR}): Every mathematician has the property of being necessarily rational but lacks the property of being necessarily bipedal.

From (1_{DR}), (2_{DR}) and (3) it follows that someone has the property of being necessarily rational but lacks the property of being necessarily rational and has the property of being necessarily bipedal but lacks the property of being necessarily bipedal, which is contradictory on two counts. However, with (1) and (2) so understood there is, according to Plantinga, "less than a ghost of a chance the essentialist will accept [them]."[7] And this is surely right, for no essentialist will hold that cyclists are essentially cyclists or that mathematicians are essentially mathematicians; why on Earth, then, would they hold that cyclists are essentially bipedal or mathematicians essentially rational? Surely had some cyclist not been a cyclist they might not have been bipedal, and likewise, mutatis mutandis, for some mathematician. What the essentialist may accept[8] are the de dicto readings of (1) and (2), below.

1_{DD}): Necessarily, every cyclist is bipedal (but it's not necessary that every cyclist is rational).

2_{DD}): Necessarily, every mathematician is rational (but it's not necessary that every mathematician is bipedal).

[7] [104, p25]

[8] Modulo discrimination against disabled cyclists.

But from these and (3) it follows only that someone is both bipedal and rational, and so trouble is avoided. So the problem only arises, in Plantinga's view, due to a de dicto/de re confusion: a de dicto reading of (1) and (2) are necessary to set up the inference as having true premisses, but it is the de re reading that is necessary in order for them to entail the troublesome conclusion.

The essentialist solutions are widely accepted; but I think that this is partly to do with a poor choice of examples on Quine's part. We are not inclined generally to think that there is any context in which it is true to say of someone that they are necessarily rational, or two-legged, and still less to think that there is any way of referring to 9 such that we could truly say of it that it could be less than 7. But I suggest that our unwillingness to think so is not due to our tracking some invariable essential properties of objects but simply due to the contingent fact that the ways in which we actually refer to 9, or to mathematical cyclists, do not evoke the relevant counterpart relations.

Something more than that may be going on in the numbers case. Even if most objects are world-bound, it may well be that numbers exist in more than one world, for it may be that they are universals. Lewis thought it was an open possibility that there were universals, and if so that they existed in more than one world. Now presumably if an object is not world-bound then what is de re possible of it should simply be what the essentialist says it is: that which is true of it in some world. That need not be to treat the truth-conditions of de re modal claims differently depending on whether or not the object to which we apply the predicate is world-bound or not; we can have a uniform story concerning the truth-conditions of de re modal claims — that a is possibly Φ iff a has a counterpart that is Φ — but simply demand that when an object is such that it exists in more than one world all and only its counterparts at a world are itself. So the counterpart theorist can have a uniform counterpart theoretic account of de re modality but agree with the essentialist as concerns the truth conditions of sentences predicating modal properties of entities that exist across worlds. Now if 9 is a universal as, for example, E.J. Lowe thinks[9], then what is possibly true of 9 is what is true of *it* — that very thing — at some world, since its counterparts at a world are just it. Now while what counts as a counterpart of an object

[9]See [84, p66].

can in general change from context to context, what is identical to an object cannot; thus we have an explanation, congenial to the counterpart theorist, for why we have a strong aversion to thinking that there is any context and any way of referring to 9 such that we can say of 9 in that context that it could be less than 7. If the counterparts of 9 are always simply itself then there is no context in which there is a counterpart of 9 that is less than 7.

However, there are far more plausible reasons to hold the inconstancy of representation de re than Quine's examples. What I find particularly motivating is that the inconstancy of representation de re lets us avoid thinking that our world is one full of coinciding entities: a conclusion many metaphysicians have been driven into by the difference in modal properties of a thing a and a thing b which share the same spacetime location. The classic example is, of course, the statue and the clay.[10] We are to suppose that at a time t some clay comes into existence, and at that exact moment is formed into a statue; then at some later time t* the clay goes out of existence, and the statue with it. The statue and the clay occupy the same area of spacetime, and so if they are distinct we have a case of coincidence: two distinct entities in exactly the same place at exactly the same times. And that the statue is distinct from the clay is, so the defenders of constancy tell us, a simple application of Leibniz's law: the clay could survive squashing but the statue couldn't, and so there is a difference in modal properties between them, and so they are not identical. The counterpart theorist of course will have no part of this, claiming that the statue is identical to the clay but that referring to it as 'the statue' typically invokes a counterpart relation different from that invoked when referring to it as 'the clay', thus explaining the difference in truth value between 'the clay could have been squashed' and 'the statue could have been squashed'. I claim avoidance of the coincidence of the statue and the clay as a benefit of the inconstancy of modality de re and, as a result, of counterpart theory.

3.1.3 Conservative and Liberal Counterpart Theory

Once we have admitted multiple counterpart relations we face a decision about just how many counterpart relations there are. In particular, we must consider the following two questions: (i) for any property Φ that is possibly had by some thing, is it the case that for every thing there

[10] Originally introduced by Gibbard [45], and which has since spawned a vast literature.

is a counterpart relation under which that thing is possibly Φ? and (ii) for any thing and any property had by that thing, is there a counterpart relation under which that thing essentially has that property? Call one who answers yes to both questions a liberal with respect to counterpart relations and one who answers no to both a conservative.[11] Very closely related to liberalism is the view that for any object x and for any set $y_1, y_2, \ldots y_n$ of possible individuals, there is a counterpart relation according to which all and only the counterparts of x are $x, y_1, y_2, \ldots y_n$[12]; a special case of this view is the claim that for every object there is a counterpart relation under which the only counterpart of that object is itself.[13] This claim, call it counterpart universalism, is stronger than liberalism, but it would be natural for the liberal to hold it.

Stronger than liberalism and independent from universalism is the doctrine that for any object x and for some property Φ of x there is a possible context in which one can truly assert that x is necessarily Φ, and that for all properties Ψ which are possibly had by some individual that for any object x there is a context in which one can truly assert that x is possibly Ψ. Call this doctrine strong liberalism. Strong liberalism presupposes liberalism, for of course one could only truly make these assertions if there are the relevant counterpart relations; but liberalism might be true and strong liberalism be false, since there might be counterpart relations that, as a matter of fact, can never be invoked in any possible context.

Later Lewis is at least halfway to becoming a strong liberal, because he thinks that for all objects there is a counterpart relation according to which that object is essentially Φ, where Φ is one of its actual properties. Furthermore, he thinks that for any object x and any property Φ of x it can be truly asserted that x is necessarily Φ by using the phrase 'x-qua-Φ is necessarily Φ'. Lewis needs to think this in order to get his truthmaking theory up and running, as we shall see in section [4.2.2]. This seems to be a departure from Lewis' earlier thinking. In [66, p87-88] he says, for example, that one cannot analyse the claim that every object could exist unaccompanied using counterparts since often the counterpart relation weighs extrinsic similarity highly.[14] Instead he

[11] This of course leaves room for a position which is neither liberal nor conservative who answers yes to one question and no to another — the Liberal Democrats of counterpart theory!

[12] Canonical counterpart theory also forces on us the constraint that y_1 to y_n must be non-actual, since it rules out distinct intra-world counterparts. One could deny this restriction and allow that the set contain actual individuals distinct from x. This would then entail that for all objects there is a counterpart relation under which every thing (unrestrictedly) is a counterpart of that object.

[13] I am assuming that every counterpart relation must be reflexive. It is a datum that every object could have been itself.

[14] See section [4.2.3].

analyses the claim as saying that for every object there could exist an unaccompanied duplicate of that object. But if liberalism is true then there is a counterpart relation according to which an object is a counterpart of x iff it is a duplicate of x, so Lewis' claim that appeal to counterparts won't help here seems to suggest that he is tacitly assuming that liberalism is false.

I am attracted to universalism (and hence liberalism), but I don't like strong liberalism. Actually, let me clarify that immediately. Strong liberalism itself comes in different strengths depending on the modality involved. A weak strong liberalism says that for any counterpart relation there is some metaphysically possible world in which that counterpart relation is invoked in some context; stronger versions would claim that the worlds in question are not merely metaphysically possible but possible in some restricted sense. I don't particularly care about the weak version but I have strong doubts about stronger versions. It seems to me that in some sense of 'can't' I simply can't invoke a counterpart relation according to which I have as a counterpart the mereological sum of a performance of Beethoven's ninth symphony, Hitler's left earlobe, and an atom in the core of the Sun. Evidence for this seems to come from our complete disinclination to hear as true any sentence saying I could have been such a thing, no matter what. There *is* such a counterpart relation, according to universalism or liberalism; but it seems entirely irrelevant to any context we may find ourselves in, and so even if there is a metaphysically possible world in which someone could truly utter (in English) 'Ross could have been the sum of a performance of Beethoven's ninth and . . . ', it seems that there is an important sense of possibility according to which it is simply not possible to truly utter this sentence (in English).

I want to hold universalism (and liberalism would have the same advantage) in part because it makes the epistemology of modality de re far more tractable. To know whether something is essentially Φ according to some counterpart relation[15] I need only know whether or not it is Φ, and so knowledge about de re necessity reduces to knowledge of actuality; and to know whether something is possibly Φ according to some counterpart relation I need only know whether it is possible that something is Φ, and so knowledge about de re possibility reduces to knowledge about de dicto possibility. Then the only problem remaining is knowing what counterpart relations are invoked in what contexts; this is simply a matter of knowing what similarity considerations are

[15]Note that what I want to know here is whether (x is essentially Φ according to some counterpart relation) and not whether, according to some counterpart relation, x is essentially Φ according to that relation.

relevant in that context, which seems a more tractable problem than discovering the essences of objects. (I will say more about the epistemological advantages of counterpart theory over essentialism below.)

Universalism also reduces the problem of the source of de re modality to the source of de dicto modality. For the de re modal facts are settled solely by what properties are possibly had and what properties could not be had, and that is settled by what states of affairs are de dicto possible. Thus while the essentialist has two problems concerning the source of modal truth — to explain what it is in virtue of which certain states of affairs could obtain and to explain what it is in virtue of which an object's properties divide into the ones it has essentially and the ones it has accidentally — the counterpart theorist only has to explain what it is in virtue of which certain states of affairs could obtain. Once she has done this the de re modal facts are there for free. Of course there is still a question as to why our utterances of sentences making modal predications express a truth in some context; but that question is a question for the philosopher of language not the metaphysician. There is no metaphysical fact to be explained here; the only thing to be explained is how context selects certain respects of similarity as relevant and others as irrelevant.

The universalist and the liberal have the advantages of there being multiple counterpart relations without the disadvantages that the strong liberal is committed to: namely that modality de re has become an 'anything goes' matter. This 'anything goes' objection to counterpart theory is considered by John Divers[16]:

> Whenever it suits your purposes that a token de re claim should come out true then all you need to do [so the objection goes] is to pick and choose, mix and match the relations judiciously so as to obtain the desired result. How, then, can deregulated [counterpart theory] embody or reflect the central norms of assent or dissent — the element of discipline over usage — that are essential to the truth-aptitude of the declarative sentences of de re modal discourse?

But no matter how damaging this objection would be against the strong liberal, it is clearly of

[16][32, p147]

no consequence to the universalist. We cannot, if strong liberalism is false, simply choose to use our language in a way as to pick out those counterpart relations that suit some purpose of ours; there is no guarantee that those counterpart relations are invocable. There *are* such counterpart relations; that much is guaranteed by liberalism and universalism; but what is not guaranteed is that we are able to invoke such counterpart relations.

3.1.4 Intra-world counterparts

As I mentioned in footnote 11, canonical counterpart theory demands that the only counterpart a thing has which is a worldmate of it is itself. I think this should be abandoned, for two reasons. Firstly, it lets us make certain appealing (to me anyway) anti-essentialist claims without admitting distinct indiscernible worlds. Consider the claim that I could have been exactly like you in fact are while you are exactly like I in fact am. If that is to come out true without appealing to intra-world counterparts we need to make appeal to a world qualitatively indistinguishable from this one but distinct from it; far more economical is to simply appeal to the counterpart relation(s) according to which you are my counterpart and I yours. (Either way we need to reject the claim that the counterpart(s) of a thing x at a world w, if there is one, is/are more like x than any other thing at w — but universalism entails the denial of this anyway.) Secondly, it just seems completely ad hoc to place a ban on intra-world counterparts. What is it about being spatio-temporally related to a thing that rules it out as being your counterpart? It just seems wholly unmotivated.

Counterpart theory combined with intra-world counterparts yields a pleasing story of how we can come to have knowledge of some claims of de re possibility. It seems to me that we can learn that something is possible for us by learning that it is true of someone who is not us but who is relevantly similar to us. Suppose I see a permanent job advertised at Oxford. I consider applying, but decide that it is pointless, as they will not be looking for someone just out of grad school. However my fellow PhD student, who started at the same time and who has the same amount of publications etc, applies and gets the job. I conclude that I should have applied, because I could have got it after all. That seems like a perfectly rational conclusion to draw to me. Furthermore, suppose my other friend is trying to comfort me and says 'Look, you probably couldn't have got

the job after all; remember that so-and-so got their undergraduate degree from Oxford.' Again, that seems to me like a perfectly rational response: to point out some relevant dissimilarity between me and the person who got the job and concluding that I was too quick to jump to the conclusion that I could have got it as well. So that seems to me to be a situation where I can be justified in claiming that something is possible for me because it is true of something which is similar to me in relevant respects; and if a relevant dissimilarity is discovered then it counts against that justification. But why should learning that something similar to me is F give me any reason to think that I *could* be F unless all it takes for it to be possible that I am F be that something similar to me is F. Of course, the possibility I learn of here is not metaphysical. I knew all along that it was metaphysically possible that I got the job at Oxford. The possibility here is some restricted de re possibility. But if similarity considerations should be relevant in this way for restricted de re possibility claims, why not for metaphysical de re possibility claims? And if similarity considerations are so relevant, that is evidence in favour of counterpart theory.

3.1.5 Contingent identity?

Does counterpart theory lead to contingent identity? Yes and no. The counterpart theorist is committed to contingent identity in the following sense: that it is true of some a and b such that a=b that a might not have been b, whereas it is never true to say this if we adopt QML with the necessity of identity. That is, counterpart theory provides counter-examples to the necessity of identity (NI)

$$\forall x \forall y (x = y \rightarrow \Box (x = y)) \qquad \text{(NI)}$$

But there is also a sense in which we don't *really* have contingent identity. After all, we never have the situation where a thing a in one world is identical to b in that world and yet there be some world in which a exists and b exists and yet are distinct. We never have that precisely because we never have that a or b exist in more than one world — individuals are world-bound. All we have is that at some world there are two distinct things which *represent* the possibility of some thing a which is actually identical to b being not identical to b at that world. There is

contingent identity, then, in the sense that a has the property of being identical to b but also has the modal property that it might not have been identical to b. In that sense it is possible that identical things be distinct. But it does not carry with it the supposedly objectionable metaphysics that contingent identity brings; it is not true to say that in some world there are distinct things which exist in our world and which are actually identical.

A useful analogy is with stage theory and temporary identity. Stage theory admits temporary identity in the following sense: that a thing a can, at time t, have the property of being identical to b, but also have the property of being such that it will not be identical to b. In that sense things which are identical now might fail to be identical in the future. So the stage theorist is committed to temporary identity in the sense that it provides counter-examples to what we will call the eternality of identity (EI).

$$\forall x \forall y (\exists t (x = y)_t \rightarrow \forall t (x = y)_t) \qquad \text{(EI)}$$

Where $(x=y)_t$ is understood as 'x is identical to y at time t'.

But the stage theorist does not think that there is some a and b which exists at a time t and are identical at t and exist at a time t* and are distinct at t*. That is what the hard-core temporary identity people — like Andre Gallois[17] — think. But the stage theorist denies that precisely because they deny that things exist at more than one time. More precisely, the stage theorist denies that the things which count as counter-examples to the eternality of identity exist at more than one time. Suppose, for example, that later in my life my body will be destroyed and my brain state transferred into a machine. Do I survive this? Well if I do, and if the eternality of identity is true, then it must not be true now that I am my body, since it will definitely not be true *then* that I am identical to what is now my body. The temporary identity theorist, however, can admit both that I am (now) identical to my body, and that when I undergo the brain transfer I will survive, after which I will no longer be identical to the body I am now identical to.

However, while the stage theorist thinks that there are cases of temporary identity, they agree with the eternality of identity theorist on the ontology of the world. They are not committed to

[17] [44]

the uncomfortable metaphysical picture a temporary identity theorist like Gallois is committed to. Gallois thinks that there are objects a and b which exist at a time t and a time t* and are identical at t and not identical at t*. The stage theorist does not think this. I do not exist after the brain transfer. I am a time bound individual according to the stage theorist; persons are instantaneous things — they are what the perdurantist thinks are merely temporal parts of persons. However, I *will* exist after the brain transfer, for what it takes for it to be true that I will exist is not that I exist at that future time but that something exists at that time which is a personal counterpart of me. And a thing need not be identical to (a future counterpart of) the body that I am now identical to in order to be a personal counterpart of me; and this explains how it can be that I will not be my body even if I now am.

So there is a difference between one's theory falsifying instances of the eternality of identity and one's theory being committed to genuine temporary identity. Likewise there is a difference between falsifying instances of the necessity of identity and one's theory being committed to genuine contingent identity. Counterpart theory does the former, but not the latter. Instances of the necessity of identity are falsified because a world can *represent* some thing which is actually identical to b as not being identical to b. But there is no *real* contingent identity, in the sense for example as Allan Gibbard defends it[18]: there are no instances of some thing a which is identical to some thing b being not identical to b at a world; for a and b only exist at one world, and at that world they are identical. Gibbard is to the counterpart theorist as Gallois is to the stage theorist; they make many similar claims concerning the temporal or modal properties of things, but the underlying metaphysics is very different indeed.

I want to claim this as an advantage for counterpart theory, although I admit it is not an advantage that will be recognised as such by many of its opponents. The advantage, to my mind, is that we secure the intuitive thought that some true identity statements might have been false without committing ourselves to a metaphysical picture that I think is rightly viewed with suspicion. The reason many of the opponents of counterpart theory will not even grant this as a benefit is that they will not consider the claim that true identity statements could have been false a claim worth trying to save. But there is, I think, a lot of intuitive force behind the

[18][45]

idea of contingent identity. Here are some statements that, as pre-theoretically as I can get in these matters, I find intuitively compelling. I am my body, but I might have been disembodied. The statue is the clay, but that very statue could have been made from bronze (even if the clay couldn't have been bronze). This instance of pain is this instance of c-fibre stimulation (or some brain-state), but any instance of c-fibre stimulation might not have been painful, and yet pains are essentially painful, hence the pain might not have been the c-fibre stimulation.[19] Of course, that each of these claims is worth saving is controversial, and each of them could take up a thesis in themselves; all I am doing is flagging the sorts of claims that might lead one, and do lead me, to find contingent identity attractive.

If I am to accept contingent identity I must say something about the Barcan/Kripke proof of the necessity of identity. The proof uses the necessity of the claim that everything is self identical, and the intersubstitutivity of co-referential expressions, to derive the necessity of identity. Essentially the argument is as follows. Necessarily, a is identical to a; that is, a has the property of being necessarily identical to itself. If a is (in fact) identical to b then a and b have exactly the same properties, by Leibniz's law. So if a is identical to b then b has the property of being necessarily identical to a. So necessarily b is identical to a if it in fact is. The proof can be formalised in different ways, but similar considerations arise in each case. Here is one formalisation in the style of Kripke.[20]

(1)	$a = b$	Ass
(2)	$\forall x \Box(x = x)$	
(3)	$\Box(a = a)$	2, $\forall$E
(4)	$\Box(a = b)$	1,3, LL/substitutivity
(5)	$a = b \rightarrow \Box(a = b)$	1,4 $\rightarrow$I

This proof has been thought questionable for even the trans-world identity theorist. One worry concerns (2); Kripke says of (2) that "every object is surely necessarily self-identical".[21] But in fact that is not wholly obvious. Here is what *does* seem uncontroversial: that necessarily every object is self-identical. But that is $\Box\forall x(x = x)$, not $\forall x\Box(x = x)$. Why should the contingent identity theorist accept the latter claim? The latter is obtainable from the former by the converse

[19] This is the position held by Smart [138].
[20] [59, p163]
[21] [ibid.]

Barcan formula (CBF) — ($\Box\forall x\Phi x \rightarrow \forall x\Box\Phi x$)s — but it is not obvious that we should accept that. CBF says that if it is necessary that every thing is Φ then every thing is necessarily Φ. But some things are contingent existents: there are worlds in which they don't exist; and in the worlds in which a thing does not exist it is not Φ (at least for many Φ). In that case couldn't it be the case that at every world everything at that world was Φ, but some actual thing not be Φ at some world because it fails to exist at that world? Well I think that could be true, but in fact I don't think this is reason to deny CBF; what is needed is to interpret the necessity in question 'weakly', so that it demands not that the proposition is true in every world but merely that it is false in no world. A thing can be necessarily F even if it is a contingent existent, provided the necessity here is weak; it is not true in every world that the thing is F, for it is not F in the worlds in which it does not exist; but it is not false that the thing is F in the worlds in which it does not exist either (for it is not not-F); I claim that in the worlds in which the thing does not exist it is neither true nor false that a is F. As such, it is weakly necessary that a is F if a is F in all the worlds in which a exists. Provided the necessity in the converse Barcan formula is interpreted weakly the formula is valid, and so we can use it to prove that identity statements are themselves weakly necessary.

This is presumably all that is wanted — Kripke wants to claim that true identity claims are true in every world in which the things exist; he does not want to hold that the things must be necessary existents.

So if resistance to the proof is to come, it must come from elsewhere. I have heard a worry about getting step (4) from (3) from philosophers who doubt the legitimacy of substituting co-referential terms into modal contexts. I won't say too much about this, but I think the onus is on them to show that this is a problem. After all, these are not people who abstain from de re modalising; they admit the intelligibility of quantifying into modal contexts. Why then should there be a problem about substituting co-referential terms into such contexts? A similar point can be made against a different version of the proof that relies on lambda abstraction and Leibniz's law (the indiscernibility of identicals) as opposed to a substitution rule. That proof runs as follows.

(1)	$a = b$	Ass
(2)	$a = b \rightarrow (\lambda z(\Box(a = z)a) \rightarrow \lambda z(\Box(a = z)b))$	LL
(3)	$\lambda z(\Box(a = z)a) \rightarrow \lambda z(\Box(a = z)b)$	1,2, $\rightarrow$E
(4)	$\forall x \Box(x = x)$	
(5)	$\Box(a = a)$	
(6)	$\lambda z(\Box(a = z)a)$	5, λ abstraction
(7)	$\lambda z(\Box(a = z)b)$	3,6, $\rightarrow$E
(8)	$\Box(a = b)$	7
(9)	$a = b \rightarrow \Box(a = b)$	1,8 $\rightarrow$I

In this proof we use the indisputable fact that identical objects do not differ in their properties and the claim that every object is necessarily self-identical to show that if a is identical to b, then since a has the property of being necessarily identical to a, b also has the property of being necessarily identical to a; from which it surely follows that necessarily a is identical to b.

Lowe[22] has expressed a worry in moving from (5) to (6). (6) attributes to a the property of being necessarily identical to a, but perhaps all that follows from (5) is that a has the property of being necessarily self-identical. If so, all that could be proven would be that b has the property of being necessarily self-identical, not the necessity of identity. One cannot move straightforwardly from 'a is necessarily self-identical' to 'a is necessarily identical to a' because the properties are obviously not the same. Everything has the property of being necessarily self-identical, but at most one thing has the property of being necessarily identical to a. So the proof of the necessity of identity is only valid, it seems, if $\lambda z(\Box(z = z)a)$ entails $\lambda z(\Box(a = z)a)$, i.e. if 'necessarily a is self-identical' entails 'necessarily a is identical to a'. But this doesn't seem too great a worry; why should this move be any worse that the move from 'a is self-identical' to 'a is identical to a', which is used in the standard proof of the *symmetry* of identity from the reflexivity of identity and Leibniz's law?[23]

I don't think the move does look any worse in the modal case, so I grant that the proof is okay for the trans-world identity theorist, and that they should accept the necessity of identity. However, there are avenues open to the counterpart theorist to deny the move from 'necessarily

[22] [81]

[23] This point is made at some length by Bob Hale in his [49].

a is self-identical' to 'necessarily a is identical to a' that are not open to the trans-world identity theorist. When the counterpart theorist says that necessarily a is self-identical, this gets translated as 'every counterpart of a is self-identical'. That, the counterpart theorist will admit, is trivially true, since they admit that it is necessary that everything is self-identical. But the claim that a is necessarily identical to a gets translated as 'every counterpart of a is identical to a counterpart of a'. And that claim is *not* uncontroversial, because the counterpart relation that is invoked in each case might not be the same. There might be a counterpart of a under some counterpart relation that is not a counterpart of a according to some other counterpart relation, and hence 'a is necessarily identical to a' can fail even though 'a is necessarily self-identical' is trivially true. What is important is that the counterpart theorist has a principled reason not to accept this move which does not rule out the legitimacy of the move from 'a is self-identical' to 'a is identical to a'; the reason the former fails is because the interaction of the identity sign with the modal operators results not in identity but in the counterpart relation, and no such problem arises in the proof of the symmetry of identity as it is non-modal; thus the counterpart theorist can accept the proof of the symmetry of identity but deny the proof of the necessity of identity.

Similar considerations show why the counterpart theorist but not the trans-world identity theorist can resist the Kripke-style proof. The counterpart theorist has a principled reason to resist substituting co-referential terms into modal contexts. The reason is that the truth of statements of the form 'A is essentially Φ' are context-dependent, and often using a different name for a thing is enough to trigger a shift in context. If the sentence 'Necessarily A is Ψ' expresses a truth, and if A=B, then *provided that there is no shift in context* it follows that the sentence 'Necessarily B is Ψ' expresses a truth. In practice, however, uttering the sentence 'Necessarily B is Ψ' is all it takes to trigger a shift in context, so that the sentence is no longer guaranteed to express a truth. So from the fact that necessarily a is identical to a and the fact that a is identical to b one can indeed infer, if there is no change in context, that necessarily b is identical to a. Usually, however, referring to a as 'b' is all it takes to trigger a shift in context in which a different counterpart relation becomes salient; and while according to the counterpart relation invoked in the previous context the thing a (=b) is essentially identical to a, according to the counterpart relation invoked in this new context the thing b (=a) is *not* essentially identical

to a.

So trans-world identity theorists are, I think, committed to the necessity of identity, and the counterpart theorist is not; and this I claim as an advantage for the counterpart theorist. Note that it is simply a special case of the advantage I claimed of counterpart theory above; namely that it lets us resist arguments that use Leibniz' law to argue for claims of non-identity on the basis of different modal properties. It is agreed on all sides that the statue has different modal properties from the clay; what counterpart theory does is let us resist the move from this observation to their distinctness. Likewise, it is agreed on all sides that some c-fibre stimulation might not have been pain or that I might have been disembodied[24]; what counterpart theory lets us do is resist the move from these observations and the further agreed truths that pains are essentially painful and bodies are essentially embodied to the conclusions that the pain is distinct from the c-fibre firing and that I am distinct from my body.

3.1.6 Inconstancy without Counterparts

I have argued for the inconstancy of modality de re, and I take this to strongly favour counterpart theory — which is a deflationary account of the essential/accidental properties distinction — since counterpart theory gives the best account of why modality de re is inconstant. It is not that someone who believes in a genuine ontological distinction between essential and accidental properties *cannot* account for the inconstancy of modality de re; it is just that I think counterpart theory offers a more plausible story than the ontic alternatives.

Laurie Paul, for example, believes in an ontological distinction between the essential and accidental properties and reconciles this with the inconstancy of modality de re by an appeal to the problem of the many.[25] We are familiar with the problem that, given certain claims about composition and decomposition, there are many distinct (but not wholly distinct) sums of atoms which have equal claim to be the referent of some name. When I use some name 'O' it is vague which of these sums I pick out. Now some of these sums may be essentially F and some merely

[24] When I say that these claims are agreed on all sides I do not mean they are not controversial; I mean only that the disagreement cuts across the counterpart theory/trans-world identity dispute. Both (some) counterpart theorists and (some) trans-world identity theorists agree that I might have been disembodied and that an instance of c-fibre firing might not have been painful.

[25] [100]

accidentally F (there being an ontological difference between the sums to account for this), and for Paul whenever I truly utter 'O could have failed to be F' in one context and truly utter 'O is essentially F' in another that is because the referent of 'O' has shifted from context to context. In this first context it picked out one of the O-candidates that, as a matter of determinate and invariant objective fact, has F as a merely accidental property, whereas in the second context 'O' picked out a distinct O-candidate that does, as a matter of objective and invariant fact, have F essentially. But I don't think Paul's account of inconstancy is satisfying. Consider the statue and the clay case. According to Paul when we consider the thing in the art gallery as a statue and when we consider it as the clay we must on each occasion be thinking about distinct (albeit largely overlapping) objects. There must be something that causes me to refer to a different statue candidate when I talk about 'the statue' than when I talk about 'the clay'. What is the cause of that, given that I think I am talking about the same thing each time? Indeed, it seems that I can simply *stipulate* that I am talking about the same statue candidate each time, and the difference in essentialist intuitions does not disappear: it still seems that the statue could lose some parts and survive whereas the clay couldn't. The most plausible story, I think, is that all that has changed is what standards of similarity are in place to determine what are the counterparts of the staute/clay.

That is my case for counterpart theory, let me now try and answer some objections. I will look at an objection to the inconstancy of modality de re (and hence to my motivation for accepting counterpart theory), but first I will consider an objection to the counterpart theoretic treatment of de re modality (independently of whether or not it treats it as inconstant).

3.2 The Humphrey Objection

A common objection to counterpart theory originates with Kripke. Kripke thought it wrong to think that I *could* be some way I am not in virtue of the fact that something which is *not* me is that way, no matter how similar that other thing is to me. He says[26]

> [I]f we say 'Humphrey might have won the election (if only he had done such-and-

[26][60, p45, fn13]

> such), we are not talking about something that might have happened to *Humphrey* but to someone else, a "counterpart". Probably, however, Humphrey could not care less whether someone *else*, no matter how much resembling him, would have been victorious in another possible world.

There are two distinct criticisms here. There is the concern Kripke raises about whether someone who is not Humphrey winning the election in some world does not account for the truth of Humphrey's possibly winning the election, and there is a concern, expressed in the subsequent sentence, concerning Humphrey's attitudes, in particular his lack of concern, concerning his counterparts.

As concerns the second criticism, there are two ways to take the objection: there is the objection that Humphrey *doesn't* care about what happens to his counterparts and the objection that he *shouldn't* care. The first objection would be unworrying. The fact that Humphrey does not in fact care about what happens to his counterparts shows nothing about the truth or falsity of counterpart theory, for of course counterpart theory could be true and Humphrey be ignorant of that fact. If Humphrey does not recognise the truth of counterpart theory then it is no surprise that he expresses no (selfish) concern for his counterparts. Analyses needn't be obviously true; that a is possibly Φ just in case a has a counterpart who is Φ need not be recognisable as true by any competent de re modaliser even if it is in fact true. So it is not clear exactly what Humphrey's lack of concern about what is true of his counterpart is meant to show; that could easily be explained by Humphrey's lack of knowledge of the analysis of de re modality. What of whether or not Humphrey *should* care about what happens to his counterparts? Well, if counterpart theory is right then he should care about what happens to them given that he cares about what could have happened to him. In that case, the objection concerning care for counterparts presupposes the objection concerning the truth-conditions of de re modal claims: Kripke only gets the result that we shouldn't care about what happens to our counterparts if he can show that what happens to our counterparts has nothing to do with what could happen to us. For if my having a counterpart who is Φ *does* account for the truth of my possibly being Φ then of course I should care about the fact that I have such a counterpart, given that I care about what could have happened to me. To put it perhaps more strongly: why should I not care about

what happens to my counterparts if it is true that *I could have been them*? A counterpart is someone who is not me, but who I could have been[27], and would have been had certain conditions been actual. Insofar as I should care about what could happen to me, then, I should also care about what happens to my counterparts.

Perhaps Kripke is right that I shouldn't care (from a selfish perspective at least) about what happens to someone else *merely* because they are similar to me; but it doesn't follow that I shouldn't care about what happens to someone who is not me when the way that person is is a way I could have been. Humphrey shouldn't feel frustrated at his not winning just because there is someone who is like him who won; but he surely has cause for concern if someone who is the way he would have been had things gone a certain way won. Kripke, of course, will deny that Humphrey could have been this person; but then that is his first objection — that a counterpart of me being Φ does not account for me being possibly Φ — not his second.

Consider an analogous objection to stage theory. According to stage theory persons are instantaneous; they do not extend throughout time. In that case, when I selfishly hope that I have a comfortable retirement, I am not literally hoping that it is *me* who has a comfortable retirement but instead selfishly hoping that a future temporal counterpart of me has a comfortable retirement. But doesn't that conflict with the fact that my hope is *selfish*? No, it does not. Because what is true of my future temporal counterpart is something that *will* be true of me. Since I can be selfishly concerned about what will happen to me, it is perfectly understandable that I can be selfishly concerned about what happens to a future temporal counterpart of me. Of course I should be concerned about what happens to my future temporal counterparts given that *I will be them*.[28] Of course, one can deny that my having a future temporal counterpart who is Φ accounts for it being the case that I will be Φ; but that is just to beg the question against the stage theorist. There is no particular argument against stage theory concerning care for temporal counterparts, just as there is no particular argument against counterpart theory concerning care for worldly counterparts; both objections presuppose that the theory in question gives the wrong truth conditions for tensed/modal language respectively.

[27] I could have had the property of being identical to C, where C is one of my counterparts, since there is a counterpart of me, C, who *does* have that property.

[28] C.f. [134]

Let us turn to Kripke's initial criticism then: that "we are not talking about something that might have happened to Humphrey". His point, I take it, is that in saying that there is a counterpart of Humphrey, something which is not identical to Humphrey, which is Φ we do not say anything about what *Humphrey* could be like. Peter Forrest makes a similar criticism; he says[29]

> On Lewis' [counterpart] theory no particular is in more than one world. As a consequence it is *not literally possible* for, say, a great uncle of Genghis Khan to be knitting champion of Ulaan Bator. On Lewis' theory, all that is possible is for a great-uncle of a counterpart of Genghis Khan to be knitting champion of Ulaan Bator.

These criticisms, however, simply beg the question against Lewis. According to Lewis, counterpart theory *does* claim that it is literally possible that a great uncle of Genghis Khan became knitting champion of Ulaan Bator, because what it *is* for that to be literally possible is for there to be great uncle of a counterpart of Genghis Khan who is the knitting champion of Ulaan Bator. Likewise, we *are* talking about something that could have happened *to Humphrey*, precisely because the truth conditions for something being possible concerning Humphrey are, according to the counterpart theorist, that it is true of some counterpart of Humphrey. The Humphrey objection, and Forrest's variant, assumes that in order for something to be possible concerning an individual a there has to be a possible world in which the very thing a exists and is that way. But obviously this is precisely what the counterpart theorist rejects. According to the counterpart theorist every individual is as it actually is in all the worlds in which it exists, precisely because it only exists in one world.[30] That will get you to the conclusion that, literally, individuals have all their properties essentially only if you think that it is a necessary condition to be possibly Φ that you are Φ in one of the worlds in which you exist. Since this is exactly what the counterpart theorist denies, the objection pretty obviously begs the question.

A similar objection occurs in Nathan Salmon's discussion of his 4-worlds paradox.[31] As we will see, Salmon thinks that the counterpart theorist simply means something different from him

[29] [42, p23, my emphasis]
[30] More precisely, every individual that wholly exists in one world wholly exists in only one world. Individuals can have parts in more than one world, given unrestricted composition.
[31] [132, p230-240]

by de re modal locutions. In that case he will agree with Kripke and Forrest that, as he (Salmon) uses the term, it is not literally true according to counterpart theory that things could have been otherwise. The 4-worlds paradox is set up as follows. Suppose at the actual world there is a ship S built of a hundred planks of wood: $\{P_1, P_2, \ldots, P_{97}, P_{98}, P_{99}, P_{100}\}$. Suppose further that the essentiality of origin is true of this ship, but that its essence admits of some variation in origin; that is, while it could have been made from a somewhat different set of planks of wood, it could not have been made from a completely different set of planks. Suppose finally that the threshold for possible change is 2%; that is, that in order for S to exist it must be built from at least 98 of the same planks of wood as it is actually built from.[32] Now consider the world W1 possible from @ in which a ship S* is made from planks $\{P_1, P_2, \ldots, P_{97}, P_{101}, P_{102}, P_{103}\}$. S* must be distinct from S, for its origin varies too much from S's origin. Now consider another world possible from @, W2, at which S is made from planks $\{P_1, P_2, \ldots, P_{97}, P_{98}, P_{102}, P_{103}\}$; this is possible because S's origin only differs by two planks from S's origin in @. Note now that the planks which S* is built from in W2 overlap enough with the planks S is built from in W1 to allow that those distinct ships could have been built from the exact same planks; for there is a world W3 possible with respect to W1 in which S* is built from the exact planks that S is built from in W2, namely $\{P_1, P_2, \ldots, P_{97}, P_{98}, P_{102}, P_{103}\}$. So we have two worlds, W2 and W3, in which two distinct ships are built from the exact same planks of wood. Furthermore, all we have been changing between worlds is what planks the ship is built from, and what is entailed by that, so we can suppose that there is no qualitative difference between worlds W2 and W3: the *only* difference between them is that the ship in W2 is S and the ship in W3 is S*. But that is an unattractive conclusion to have reached; as Salmon says "How can these two ships, having the very same original matter and structure, not be one and the same ship? After all, it would seem that a ship is nothing *over and above* its parts put together in a certain way."[33] The worry is that we seem committed to mere haecceitistic difference between the two worlds W2 and W3; and the thought that two worlds can be qualitatively identical differing merely haecceitistically is not a comfortable one. In response to this problem Salmon denies that W3 is possible even though he admits that it is possibly possible; that is, while W3 is possible with

[32] The numbers are chosen for convenience and are irrelevant to the argument.
[33] [ibid. p232]

respect to W1 and W1 with respect to @, W3 is not possible with respect to @. This lets him avoid the result that two possible worlds can differ merely in haecceitistic properties. I prefer a counterpart theoretic solution to Salmon's 4-world paradox: for the counterpart theorist, W2 is simply *wholly* indiscernible from W3: there is not even a haecceitistic difference; indeed, the counterpart theorist would probably simply *identify* W2 and W3. On this solution the ship in W2(=W3), call it S**, is a counterpart of S in @ and a counterpart of S* in W1; but while S** is a counterpart of S* and S* a counterpart of S, S** is not a counterpart of S, since its origin is too different.

Salmon anticipates this response; he says[34]

> [A] common reaction to the Four Worlds Paradox is to resist its conclusion by holding that the worlds W2 and W3 are really one and the same possible world, and that we are simply calling a single ship by two names, 'S' and 'S*'. . . The basic idea is that . . . [s]trictly speaking, it is not true that there is a possible world in which the ship that we have called 'S' in @ — that very entity — is originally constructed from any matter (planks) other than the very same matter (planks) from which it is actually constructed in @. . . Rather . . . they [the ships in W1, and W2/W3] are proxies or 'counterparts' of one another.

And he rejects this response on the grounds that it does no justice to the original intuition motivating the example: namely, that some slight variation in the origin of the ship is possible. He continues[35]

> Whatever particular form this account takes, it flies in the face of a straightforward, literal construal of the initial, plausible assumption that some artifact is such that it might have been originally made of slightly different matter. . . [C]ounterpart theory . . . is, at bottom, just a particularly inflexible brand of essentialism. By denying that there is a possible state of affairs in which the very ship S from @ is made from ever so slightly different matter, the counterpart theorist holds, even if

[34] ibid. p232-233. I have changed the names of Salmon's worlds and ships, and so have altered them throughout this quote.
[35] [ibid. p234-236]

> tacitly, that in a strict sense it is absolutely impossible for the very ship S, rather than some other ship, to have been made from matter even only an atom different. The counterpart theorist can mouth the words 'Ship S might have been made from slightly different matter', but any such pronouncement by the counterpart theorist . . . is a verbal camouflage that merely postpones the inevitable. What matters is what the counterpart *means* by those words, and more importantly, what is *not meant* by those words.

Salmon is rejecting the counterpart theoretic solution to the 4-worlds paradox because instead of resolving the puzzle generated by the supposition that the ship could have had a slightly different origin, the counterpart theorist denies this supposition by holding that while there are *proxies* of the ship that originate slightly differently, *that very ship* S has its origin essentially, down to the very last detail. I don't think this response is good.

Salmon's criticism is that the counterpart theorist *means* something different from the trans-world identity theorist when they say that a could have been Φ. He said "What matters is what the counterpart theorist *means* by those words, and more importantly, what is *not meant* by those words."[36] The implication being, I take it, that the counterpart theorist does not mean that the very thing a could have been Φ when they utter the English sentence 'A could have been Φ' but rather they mean that some replica of a is Φ. But I do not see the counterpart theorist as disagreeing with the trans-world identity theorist about the *meaning* of de re modal language; both the counterpart theorist and the trans-world identity theorist are speaking English — their debate is held using the English words 'could' and 'must' — and their disagreement is over the *truth conditions* of propositions expressed using such language. One can object that either party has not specified the truth-conditions of such propositions accurately, but it is a mistake and unhelpful for one side to accuse the other of changing the subject by meaning something different by the terms involved. It seems clear that there is a disagreement between the counterpart theorist and the trans-world identity theorist, and if there is a disagreement then this shows that they do *not* mean something different by de re modal language, for if they did then they would be talking past one another, not disagreeing; the fact that there is genuine

[36] Salmon op cit. p236.

disagreement shows that they mean the same thing but disagree over the truth conditions of the propositions expressed.

Consider again an analogy in the debate on persistence, in particular the debate between those who think that (most everyday) objects *endure* through time — i.e. are wholly present at each moment they exist, and those who think that objects *perdure* through time — i.e. have a temporal part located at each moment they exist, which exists at no other moment. Isn't the right way to view this debate as a debate over what it is to persist? — over the truth conditions of claims of the form 'A existed from time t until time t*'? Isn't it a complete mischaracterisation of the debate to describe it in such a way that the endurantist and the perdurantist mean something different by 'persists'? If that was a correct characterisation then, once each side was aware of the other side's position, there would be no debate over whether or not things persist by enduring or perduring. For the endurantist it would simply be a tautology that if anything persists it does so by enduring, and that nothing persists by perduring; likewise mutatis mutandis for the perdurantist. The endurantist would grant that anything that persisted *as the perdurantist uses the term* would do so by having a temporal part at every moment it exists, and what they would deny was that things persist in the perdurantist's sense; likewise mutatis mutandis for the perdurantist. But that is not how the debate in fact goes: both sides in fact agree that things persist; and the disagreement is over whether or not they do so by perduring or by enduring. So the meaning of 'persistence' must be agreed upon between the two sides; the disagreement is not over the meaning of 'persists' but over *what it is* for a thing to persist. Likewise, if the counterpart theorist meant something different by de re modal vocabulary it would not be common ground between them and the essentialist that some things could have been other than they are. But this *is* common ground; where the disagreement comes is over *what it is* for something to be possibly different from how it in fact is. Another way of putting what is essentially the same point is this. We all use the English terms 'could' and 'persists'; but our usage of 'could' does not settle the issue of whether things could have been otherwise in virtue of having a counterpart who differs in some way or in virtue of being different in some other world, nor does our usage of 'persists' settle whether or not things have temporal parts. To settle these questions we need to do more than reflect on the meaning of the terms involved: we need to do some metaphysics. It is an

open question whether or not things have temporal parts, but it is not an open question whether or not things persist: that things persist is a datum of the debate; the question is *how* do they persist? Likewise, the counterpart theorist and the essentialist can agree that some things could have had different properties; the question is *what are the truth-conditions* of such a claim? Of course, Salmon can use his words as he likes, and so a fortiori he can use de re modal language in such a way as to make counterpart theory analytically false as a theory concerning the truth conditions of that language. But then of course it is he who has simply changed the subject. Salmon presumably thinks he is speaking English when he uses de re modal locutions; but the counterpart theorist has just as much claim to be doing so as he.

I am accusing Kripke et al of simply begging the question against the counterpart theorist when they say that according to the counterpart theorist things couldn't literally be other than they are; but question begging arguments can still have some dialectical force, if there is strong intuitive ground to the position of the question-begger.[37] But even putting the question begging worry aside, there is something extremely odd about Kripke's and his supporters' criticisms. What is their demand? That for an individual a to be possibly Φ there must be another possible world in which, literally, a exists and is Φ? It cannot be that. For Forrest possible worlds are (possibly uninstantiated) structural universals — individuals are not literally parts of any of them: their only parts are simpler universals. And Kripke also is no realist about possibilia. The only person who believes in the literal identity of things through possible worlds is the modal realist who allows overlap. There are two ways one could have such a theory: by thinking of ordinary individuals as being modal-continuants — mereological sums of world bound individuals — or by thinking of ordinary individuals as being wholly located in each Lewisian world. Lewis' counterpart theorist treats identity across worlds as the stage theorist treats identity across time — by denying it, and replacing it with a less strict similarity relation. The modal-continuant theorist treats identity across worlds like the perdurantist treats identity across time — a thing exists in multiple worlds by having different worldly-parts at every world in which it exists (analogous to having different temporal parts at each moment of time); and just as being Φ at a time amounts to having a temporal part that is Φ according to the perdurantist, so will

[37] Consider for example the dialectical situation against the dialetheist. Many are unconvinced, to say the least, that there are true contradictions; but to argue for that without assuming the law of non-contradiction is not easy.

being possibly Φ amount to having a worldly-part that is Φ according to the modal-continuant theorist. The other option is to treat identity across worlds like the endurantist treats identity across time — by allowing that individuals are wholly located at every world in which they exist just as the endurantist thinks individuals are wholly located at every time at which they exist; an individual will be possibly Φ, then, just in case it is Φ at one of the worlds in which it wholly exists.[38]

The only people that can claim to have literal identity across worlds are those that combine Lewisian realism with modal-continuant theory or whole location in each world. The ersatzer on this point sides with the Lewisian counterpart theorist in saying that there is no *literal* trans-world identity but merely that other worlds *represent* things as existing there. It is *how* other worlds represent this that is up for debate; but all parties agree that it is a mere representation. Lewis makes much the same point.[39] He asks "[D]oes it ever happen that something which exists according to one world and something which exists according to another are identical?"[40] Yes, he answers; and this is the answer all parties should agree on, no matter what they think about either the nature of worlds or about counterpart theory. He says "[C]ounterpart theorists and ersatzers are in perfect agreement that there are other worlds (genuine or ersatz) *according to which* Humphrey — he himself . . . — wins the election. And we are in equal agreement that Humphrey — he himself — is not *part* of these other worlds."[41] Where there will be disagreement is not on this question but on the further question "What is representation de re? How does a world, genuine [Lewisian] or ersatz, represent, concerning [some individual], that he exists?"[42] It is possible that I should have had blonde hair, so everyone who believes in worlds, no matter what they think about their nature, should agree that there is a world according to which I have blonde hair. Disagreement will arise only in how it is that a world represents my having blonde hair. The Lewisian realist with overlap will say that there is a Lewisian world in which I exist and have blonde hair; Lewis would say there is a Lewisian world in which a counterpart of me exists and has blonde hair; one type of linguistic ersatzer will say that there is a maximally consistent set of sentences one of which is 'Ross has blonde hair'; while another will say that

[38] See [87]
[39] [66, p192-198]
[40] [ibid. p194]
[41] [ibid. p196]
[42] [ibid. p194]

there is a maximally consistent set of sentences one of which is 'a counterpart of Ross has blonde hair'. So it is not at all clear what the Humphrey objection is meant to be. If the objection relies on the fact that the counterpart theorist represents a world as containing an individual by a means other than strict identity between the individual and some thing which is literally a part of that world, then the objection applies to erstazist theories as much as to counterpart theory.

3.3 Kripke and the Essentiality of Origin

Richard Cartwright makes the following complaint against essentialism.[43]

> Chief among [the perplexities of essentialism] is the obscurity of the grounds on which ratings of attributes as essential or accidental are to be made. Apparently, in any particular case, one is simply to reflect on the question whether the object in question could or could not have lacked the attribute in question. . . But the criteria to which one appeals in such reflection are sufficiently obscure to leave me, at least, with an embarrassingly large number of undecided cases. . . [T]he distinction is [thus shown to be] a good deal less clear than essentialists are wont to suppose.

Counterpart theory lets us engage in essentialist talk while respecting Cartwright's thought. Yes, essentialist attributions leave a large number of cases undecided; yes, the criteria for a property being essential are muddled and vague; and the result of this is that there is no genuine ontological distinction between properties here. Rather, we are carving such a distinction, and we obviously have not done so with the explicitness that would lead to a settled and constant list of what properties are essential and what properties accidental.

In this section I want to take on what I think to be the biggest challenge to the deflationary way of thinking about essentialism. I want to consider Kripke's attempt to *prove* that certain properties are had essentially. If Kripke's proofs are successful, and if they start from premisses that are not themselves susceptible to a deflationary treatment, then they motivate a genuine

[43] [23, p626]

ontological distinction between essential and accidental properties, since it is obviously not a mind-dependent matter whether the result of a proof holds true: if there is a proof that p, and the premisses of the proof hold no matter what we might think about them, then p is true no matter what we might think about it. What I want to consider, then, is Kripke's 'proof' of the essentiality of origin in the infamous footnote 56 of Naming and Necessity. I will argue that Kripke's proof, and attempts to reconstruct it by Nathan Salmon and Guy Rohbraugh and Louis deRosset, fail, and hence pose no problem for the deflationist.

The essentiality of origin, as applied to persons, is a very appealing doctrine. Some philosophers deny it, of course, as is true of any interesting thesis; but most of us find it a very intuitive thesis when we encounter it, and non-philosophers have, in my experience, immediately agreed with it when confronted by it. For the counterpart theorist this constitutes good evidence that in most ordinary contexts we select a counterpart relation according to which something can only count as a counterpart of a person if it has a similar origin to that person. But Kripke aims to *prove* the essentiality of origin, and if successful this undermines the counterpart theorist's deflationary view.

Kripke's 'proof' assumes the necessity of identity. One could stop here and complain that he is begging the question since, as I argued above, part of the attraction of the deflationary counterpart theoretic view is that the necessity of identity fails. Nevertheless, in this section I will grant the necessity of identity for the sake of argument, and show that even given this Kripke's proof fails.

In footnote 56 of *Naming and Necessity*, Kripke gives the following argument for the essentiality of origin of a table from the substance from which it originates.[44]

1. Let 'B' be a name (rigid designator) of a table, let 'A' name the piece of wood from which it originally came. Let 'C' name another piece of wood.

2. Suppose B were made from A, as in the actual world, but also another table D were simultaneously made from C.

3. In this situation B$\neq$D.

[44] [60, p144 Fn.56]

4. Assumption: there is no relation between A and C which makes the possibility of making a table from one dependent on the possibility of making a table from the other.

5. Therefore: Even if D were made by itself, and no table made from A, D would not be B.

On its simplest and most literal reading the argument is flawed, as it is clearly not for the essentiality of origin at all. The conclusion of the argument is that a table D, which was introduced explicitly as being a table which was not B, would not be B either in a world in which B was made from A or in a world in which no table was made from A. But that is a claim that follows simply from the necessity of distinctness (which Kripke is assuming[45]); since D is not, *ex hypothesi*, B, necessarily (by the necessity of distinctness) B is not D. But clearly that falls far short of the essentiality of origin. For that principle to be established we do not want to know that some table which is made from C and which in fact is not B could not have been B. We want to know whether B could have been made from C. This point was well made by Nathan Salmon. He says[46]

> So far Kripke has shown only that in any possible world in which a table D is constructed, D still is not the same table as B. What he needs to show [to prove the essentiality of origin] is that in any possible world in which *a table* (meaning *any* table) is made from hunk C, *that very table* made from hunk C still is not table B.

Salmon tries to save Kripke's argument, but I will argue that the revised argument is unconvincing. I will then offer an alternative interpretation of Kripke that I think plausibly fits the text; but I will show that this argument fails due to a subtler flaw.

Let's put the problem into Salmon's notation. Let 'T(x,y)' denote that table x has been made from hunk of wood y. Salmon's problem is that although Kripke has proven E1, below, he needs to prove E2, which is the essentiality of origin proper.

$$\Box[T(D,C) \rightarrow D \neq B] \tag{E1}$$

$$\Box\forall x[T(x,C) \rightarrow x \neq B] \tag{E2}$$

[45] ibid.
[46] [131, p709]

He then seeks to show how Kripke might move from E1 to E2. His first suggestion is the principle P1[47]

$$\Diamond T(D,C) \rightarrow \Box[Dexists \rightarrow T(D,C)] \qquad \text{(P1)}$$

P1 will close the gap between E1 and E2, but it clearly begs the question as it "is tantamount to the assumption that if it is possible for a given table to originate from a certain hunk of matter, then it is necessary that [should the table exist] the given table originate from the hunk of matter."[48] Given the necessity of distinctness that just is the essentiality of origin. However, Salmon does not believe that Kripke has used this premise. Salmon claims that Kripke's missing premise is that "in any possible world, any table originating from hunk C is the very table D and no other." Or, more precisely, "If it is possible for a table x to originate from a hunk of matter y, then necessarily, any table originating from hunk y is the very table x and no other."[49] In symbols

$$\forall x \forall y[\Diamond T(x,y) \rightarrow \Box \forall z(T(z,y) \rightarrow z = x)] \qquad \text{(P2)}$$

Salmon is happy with P2. It gets us to the conclusion E2, because it "tells us that in any possible world any table originating from hunk C must be D and not B."[50] As Salmon says, the argument from P2 to E2 "uses origin as a (necessarily) sufficient condition for being *this very table* in order to prove that origin is also a (necessarily) necessary condition."[51] P2, says Salmon, is a "strong essentialist principle concerning tables and their origins."[52] But he embraces this. He claims that Kripke cannot deduce the essentiality of origin without this.

I find the argument for the essentiality of origin using P2 as convincing as that using P1: not very. Since I started out with no essentialist inclinations, appealing to essentialist intuitions will not convince me. But even if I had sympathies for essentialism, P2 doesn't look to me to be a principle I would accept. It says that there is only one possible table that can be made from any particular block of wood; but why should I accept that? Given a block of wood I could make a table that was four-legged or three-legged, tall or short, round or square, thin or wide. Am

[47] Salmon doesn't include the qualification that D must exist in the worlds in question, but this is surely needed.
[48] ibid. p.711
[49] Ibid.
[50] Ibid. p.712
[51] Ibid.
[52] Ibid.

I to believe that it would be the same table I was making in each case? Why should that be? Suppose I make table B from block A. Certainly it is possible that I make something from A which is *not* a table, call it E. Do we feel any compulsion to think that E would be B? I think not. But then why the compulsion to think that any *table* made from A would be B? If *things* can be made from A which are not B, then why could there not be *tables* made from A which are not B? I suppose Salmon would deny that I could make *anything* from A which is not B, but I see no reason to think that. It is possible that A had been fashioned into a bust of Socrates. Consider a world in which this occurred; why should we be compelled to describe this world as one in which the table B is a bust of Socrates? More intuitive, surely, is the view that in this world B was not made from A but rather the bust of Socrates was made from A. Indeed, that line of defence of P2, that there is no *thing* let alone no table that could be made from A that is not B, would rule out another essentialist principle: the claim that an object essentially belongs to the kind to which it actually belongs; for since we can obviously make A into different kinds of thing, it would follow that B can belong to different kinds. But between the principle that things belong to their kinds essentially and P2, I find the former more compelling.

Furthermore, while P2 does not beg the question it does seem to me to be too strong a premise to assume; after all, P2 will get you straight to the essentiality of origin if we add the (surely plausible) premise that it's possible to make distinct tables from distinct blocks of wood; there is no need to appeal to E1. For suppose that the essentiality of origin were false. Then there would be some table B made from a block of wood A such that B could be made from a block of wood C such that A$\neq$C. So there is a possible world, W1, in which B is made from C. Since it is possible to make distinct tables from distinct blocks of wood there is a possible world W2 in which a table D is made from C such that D$\neq$B. But if origin is a (necessarily) sufficient condition for a table being the table it is, then it follows from the fact that there is a world in which B is made from C that the table made from C in W2 is B. So the table made from C in W2 both is and is not B. That is a contradiction, so the supposition that the essentiality of origin is false would have to be abandoned. Note that this argument didn't use the necessity of distinctness, merely the necessity of the principle that two distinct things cannot, in the world in which they are distinct, also be identical. I.e. the principle $\Box(a \neq b \rightarrow \neg(a = b))$ and not the

principle $a \neq b \rightarrow \Box(a \neq b)$.

There are, I suppose, some philosophers who will agree with P2. If we talk of constitution rather than origin and hold the view that constitution is identity then, if we also hold the necessity of identity, we will hold P2 where 'T(x,y)' is read as 'x is constituted out of y'. This is perhaps Kripke's goal: to derive 'deep' essentialist truths from weaker ones such as the necessity of identity in conjunction with non-essentialist principles about material constitution. But I think that argument would be a bad one. I don't want to argue against either the claim that constitution is identity or the necessity of identity here, but I want to claim that holding both those claims is something we should not do. This is because I think the best way to make sense of constitution as identity is by adopting counterpart theory, for counterpart theory best accounts for the intuitively different de re modal properties that the one thing has when considered as constituting or when considered as constituted.[53] But if one is a counterpart theorist then one is unlikely to accept the necessity of identity.[54] In that case, the two principles needed to prove the essentiality of constitution are in tension with each other. One should not hold them both, and hence the argument for P2 is unconvincing.

I wish to suggest an alternative interpretation of Kripke's argument, which avoids appeal to P2. The interpretation aims to respect the importance Kripke places on the assumption at step (4), which I will call the independence assumption. When Kripke is making clear the presuppositions of his argument he says that we can either assume the necessity of distinctness or the necessity of identity with Brouwer's axiom $p \rightarrow \Box\Diamond p$. But no matter which of these we choose, he thinks the independence assumption is essential. He says "In any event, the argument applies only if the making of D from C does not affect the possibility of making B from A, and vice-versa."[55]

The reconstructed argument goes as follows. Suppose, as Kripke would have us suppose, that I make B from A and that nothing is made from C. Now consider a world in which B is made from A and some table is made from C. The table that is made from C in that world is not B, since tables made from distinct blocks of wood are distinct. (And we can conclude that A and C are distinct in this world since they are distinct in the actual world — this is where the necessity

[53] See [66, p252-253]

[54] For although I am my body, I might not have been, since the bodily counterpart relation is not the personal counterpart relation. See [63, p47-54].

[55] Kripke, loc cit.

of distinctness is crucial in the reconstructed argument.[56] If we only have the independence assumption and not the necessity of distinctness then that leaves it open that B could be made from C in the worlds in which A=C.) Now consider a world in which no table is made from A but a table is made from C. Still, it seems, the table that is made from C cannot be B. For if it were, then that would inhibit the possibility of making B from A since, again, tables made from distinct blocks of wood are distinct. So if I made table C from block B then I could not make C from A; but that would be to violate the independence assumption. The possibility of making a table from A should not be affected by making a table from C; since making B from C would affect the possibility of making B from A it cannot be a genuine possibility. Thus B cannot be made from C, and since we are dealing with arbitrary tables and blocks of wood the essentiality of origin for a table from its originating matter follows.

Let me make a couple of remarks on my reconstruction, mainly to stave off the accusation that I am misinterpreting Kripke and hence knocking down the argument of a straw man. My reconstruction is most plausible as a piece of Kripke exegesis if we understand the term 'D' in his argument as not being a name for a particular table but rather as a variable, or even a definite description, which denotes in a world w *whatever* table happens to be made from C in w. This reading is not without plausibility. The terms 'A', 'B' and 'C' are explicitly introduced as being names of the relevant tables or blocks of wood; but crucially this is not done with 'D'. Kripke writes[57]

> Let 'B' be a name (rigid designator) of a table, let 'A' name the piece of wood from which it actually came. Let 'C' name another piece of wood. Then suppose B were made from A, as in the actual world, but also another table D were simultaneously made from C.

These are the first occurrences of 'A', 'B', 'C' and 'D' in the argument. Why are 'A', 'B' and 'C' introduced as being names, and in quote marks, whereas 'D' is neither? One answer would be that 'D' is not intended as a name. Since we are reminded in the quote that names are rigid designators, this would suggest that 'D' is *not* a rigid designator; rather, the reference of 'D'

[56]The necessity of identity and Brouwer's axiom would do just as well, since they jointly imply the necessity of distinctness.

[57]Kripke loc cit.

is to vary from world to world depending on what is made from C in each world. In that case we cannot use the necessity of distinctness to argue that necessarily B≠D, for the necessity of distinctness is only guaranteed when the terms flanking the identity are rigid designators. Rather we would use the independence assumption to show, as we did above, that no matter what is designated by 'D', it cannot be B.

My reconstruction makes crucial use both of the independence assumption and the necessity of distinctness.[58] As such I think my reconstruction is an improvement over Salmon's reconstruction which, as we saw, uses neither the independence assumption nor the necessity of distinctness/identity. At the end of the day, however, I hope it will prove of interest to examine the reconstructed argument independently of its plausibility as a piece of Kripke exegesis.

Now if Kripke's argument as I have reconstructed it is sound then it is easily generalisable to other essentialist principles. Take, for example, the essentiality of a person originating from the sperm and egg from which they in fact originate. We could demonstrate that as follows. Consider a man A and a woman B. They have a child, C. Now consider another couple, D and E. They in fact do not have any children, but that is a contingent matter: they could have had a child. So there is a possible world, call it w, in which D and E have a child, call her F. Now whether or not one couple has children has no bearing on the child-producing abilities of another couple. So in particular the fact that D and E have had a child in w does not affect the child-producing abilities of A and B in w. It follows then that any children which A and B could produce could be produced by them in w. So in particular it could be the case that C is produced by A and B in w. Well imagine this is in fact the case (in w). The child of one couple is distinct from the child of another couple, so C≠F. Now imagine it not to be the case in w that A and B have the child C. Nevertheless, if we are to respect the principle that whether or not one couple has children has no bearing on the child-producing abilities of another couple then it still cannot be the case that F=C. For if the child of D and E was C then that would mean that if A and B had a child then it could not be C, since the children of two distinct couples are themselves distinct. So C cannot be born from D and E. D and E are an arbitrary couple, so C cannot be born from any parents other than A and B. A, B, and C are arbitrary people, so the origin of a person is

[58] Or the necessity of identity and Brouwer's axiom.

essential to that person.

Unfortunately, the argument seems to generalise *too* well. Consider the following argument. Every (concrete) object that exists occupies some space at every moment of its existence; so every object has a 4-dimensional area of spacetime which it occupies in our world. Pick one of these objects and call it A, call the 4-dimensional area of spacetime it occupies A*. Now, the domain of our world might have been larger than it actually is; that is, there might have existed objects which don't actually exist. So there is a possible world, call it W, in which an object B which doesn't actually exist exists and occupies a 4-dimensional area of spacetime B*. Whether or not one object occupies spacetime has no bearing on the spacetime-occupying abilities of another object. So in particular the fact that B occupies space in W does not affect the spacetime-occupying ability of A in W. It follows then that any area of spacetime which A can occupy, it can occupy in W. So in particular it could be the case that A occupies A* in W. Well imagine this is the case (in W). Two objects cannot occupy the same area of spacetime, so A*≠B*. Now imagine it not to be the case in W that A occupies A*. Nevertheless, if we are to respect the principle that whether or not one object occupies spacetime has no bearing on the spacetime-occupying abilities of another object then it still cannot be the case that A*=B*. For if B had occupied A* then that would mean that if A had occupied spacetime then it couldn't have occupied A*, since two objects cannot occupy the same area of spacetime. So B cannot occupy A*. A was an arbitrary object, so B cannot occupy an area of spacetime which is actually occupied by any other object. B is an arbitrary object, so no object can occupy an area of spacetime which is actually occupied by any other object. Given the further (plausible) assumption that every area of spacetime is possibly occupied by some object we get the conclusion that an object's spacetime location is essential to it.

This parody argument purports to prove the essentiality of spacetime location, a conclusion we are unlikely to accept. We have two options then. Either we must find a relevant disanalogy between the parody argument and the argument for the essentiality of origin of humans, or we must reject that latter argument. Well, there is a disanalogy between the arguments. The parody argument relies on the necessity of the principle that no two objects can share the same spacetime, and that assumption plays the same role as the assumption in the original that the

children of distinct couples are also distinct. This, one might think, puts the original argument on a firmer footing than the parody argument, for it is not clear that it is necessary that distinct objects cannot share the same area of spacetime; indeed, it's not entirely uncontroversial that it's even true.[59] But this analogy is not a relevant disanalogy. As we shall shortly see, the parody argument doesn't work even granted the assumption that as a matter of necessity distinct objects occupy distinct spacetimes; and once we see why the parody argument fails we will be able to see why the original argument also fails.

What is wrong with the parody argument is this. There is an ambiguity in the phrase 'whether or not one object occupies spacetime has no bearing on the spacetime-occupying abilities of another object'. This could mean either that the fact that some object occupies spacetime doesn't affect whether or not another object can occupy spacetime *per se*, or it could mean that the fact that some object occupies spacetime doesn't affect whether or not another object can occupy a *particular* spacetime. We can grant the truth of the former claim, but that will not establish the essentiality of spacetime location. For the parody argument to go through the latter claim must be true. But that is not true. *Of course* the fact that an object A is located in spacetime affects whether another object B occupies a *particular* spacetime, for the simple reason that B cannot have the same spacetime location as A. (Remember that we are taking for granted the necessity of the claim that distinct objects cannot occupy the same spacetime.) The parody argument claims that B can never occupy A*, as that would violate the principle that 'whether or not one object occupies spacetime has no bearing on the spacetime-occupying abilities of another object'. But it does *not* violate that principle in the sense in which the principle is true: B occupying A* does not affect whether A is spacetime located. It only violates the principle in the sense in which the principle is false: B occupying A* does affect the *particular* spacetime which A can occupy; for if A* already contains B then it cannot also contain A. So either the parody argument relies on a false principle, in which case it is unsound, or it relies on a true principle but which doesn't imply the conclusion, in which case it is invalid.

There is a similar ambiguity in the claim in the original argument that 'whether or not one couple has children has no bearing on the child-producing abilities of another couple'. This could

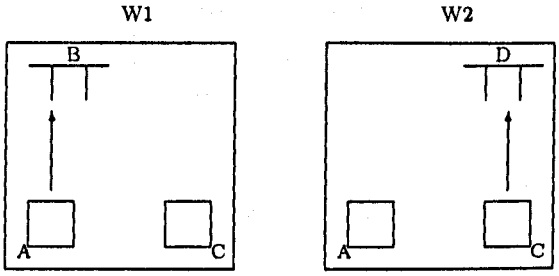
[59] Many think that the statue and the clay case is a case where distinct objects occupy the same area of spacetime.

mean that the fact that one couple has children does not affect whether or not another couple has children *per se*. We can grant that this is true, but that is not strong enough for the argument to go through. The argument claims that C cannot be born from D and E, because that would go against the child-producing abilities of A and B. But clearly C being born from D and E doesn't violate the ability of A and B to have children *per se*, they can still have children that are not C. For the argument to go through we need the stronger reading of the principle to be true, that whether or not a couple has children has no bearing on the ability of another couple to produce *particular* children. But I can make against that claim the parallel point to the one I made against the stronger reading of the claim used in the parody argument: *of course* the fact that the couple D and E had a child affects the *particular* children another couple can have, for the other couple can now not have the child that D and E had (since the children of distinct couples are distinct). In particular, if D and E have the child C then that prevents A and B from having the child C. For one to assume the truth of the stronger principle (as is necessary for the argument to be valid) we need to deny this as a possibility. But that is to deny that D and E can have C as a child, which is just to assume the essentiality of origin. So either the argument relies on a principle which is too weak for it to be valid, or the principle it relies on begs the question and assumes the essentiality of origin in order to prove it. Either way the argument is unconvincing.

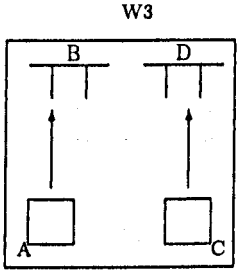
As will by now be obvious, the reconstruction of Kripke's argument will be invalid if we read the independence assumption (4) weakly, as saying that making a table from C does not affect whether a table *per se* can be made from A. We need a stronger reading that talks about particular tables. But even the stronger reading is ambiguous. Consider the strong reading of Kripke's principle: it says that whether or not a table B is made from a block of wood A in some world does not affect the *possibility* of making a particular table, D, from a block of wood C($\neq$A). But this could mean either one of two things. It could mean that if it is a possibility to make D from C then there is a world in which D is made from C *and* B is made from A. Or it could mean that if it is a possibility to make D from C then in any world in which B is made from A *it is possible in that world* that D be made from C. The distinction here is between limit possibility and compossibility. We can accept the claim that making D from C does not limit the possibility

of making B from A. That is, in any world in which D is made from C it is still possible in that world that B be made from A. But that won't get Kripke the conclusion he needs. The validity of the argument requires the compossibility claim: that since it is possible to make B from A, if it is possible to make D from C then there exists a world where both D is made from C *and* B is made from A. But there seems to be no reason to accept this claim unless we already believe in the essentiality of origin; certainly the compossibility claim $(\Diamond p \wedge \Diamond q) \rightarrow \Diamond(p \wedge q)$ is not valid, not even in S5. (After all, just let p be a contingent proposition and q its negation.) Whereas the limit possibility claim $(\Diamond p \wedge \Diamond q) \rightarrow \Diamond(p \wedge \Diamond q)$ is valid in S5 (but not B or S4).

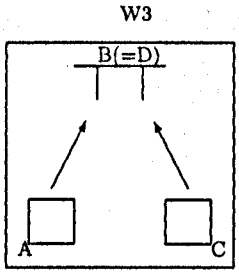
The point here is perhaps easier to see if we have a diagram. We are asked to suppose that in the actual world B is made from A and no table is made from C, but that it is possible to make a table from C. So, with W1 as the actual world and W2 as a world possible with respect to W1 (and with the arrow representing composition; so that an arrow pointing from x to y means that y is made from x) we have.



Kripke needs to be able to deduce an absurdity from the supposition that D=B. To this end he needs to show that the possibility of W1 and W2 entails the possibility of W3 with respect to W1.



W3 is an impossible world, thinks Kripke, if D=B; so if the possibility of W3 follows from the possibility of W1 and W2 then, on the assumption that D=B, we get the conclusion that W3 is both possible and impossible. This is an absurdity, so we can conclude that D≠B. In fact it is not obvious that W3 is impossible if D=B. It looks impossible due to how I have drawn it, because it looks as though we have the same physical object wholly located in distinct and non-overlapping areas of space; and that looks to be impossible. Although to be fair it's not *obviously* impossible. Josh Parsons has defended the claim that physical objects can entend — be wholly located at multiple spatial locations[60]; and Richard Cartwright famously defended the view that there are scattered objects.[61] Why then could there not be entended scattered physical objects, as B would be were W3 actual? But even granting the impossibility of entended scattered objects it is not clear why, if we are not essentialists about origin (and at this stage in the proof we must be able to consistently assume that that thesis is false), the table B(=D) cannot be made from the fusion of blocks of wood A and C. In which case W3 is better drawn as



I cannot think of a reason why W3* is impossible that does not assume the essentiality of origin. But I suppose Kripke's principle could be read as saying that the possibility of making D solely from C does not affect the possibility of making B *solely* from A, so for the sake of argument let us grant Kripke the claim that W3 is impossible if D=B. The reductio only goes through if the possibility of W1 and the possibility of W2 entails the possibility of W3. But that entailment only succeeds, it seems, if the compossibility claim is true: that if it's possible to make B from A and possible to make D from C then it is possible to make B from A *and* make D from C. But the compossibility claim does not look plausible unless one accepts the essentiality

[60] [98, p399-418]
[61] [25]

of origin. What *is* plausible is the limit possibility claim; this says that the possibility of W1 and W2 entails that there is a possible world W4 in which B is made from A and *it is possible* (in W4) that D is made from C. The possibility of W4 entails the possibility of W5, where D is made from C. But we are *not*, given only the supposition of limit possibility, entitled to conclude that in W5 B is also made from A; that would be to confuse limit possibility with compossibility. Indeed, if we assume that B=D and that W3 is an impossible world then we will conclude, and will be perfectly entitled to conclude, that it is definitely *not* the case in W5 that B is also made from A.

Let us sum up our objections to the reconstruction of Kripke's argument. The premise he relies on is ambiguous; it can be read in (at least) three ways:

- The weak reading: making a table from C cannot affect whether a table *per se* can be made from A.
- The medium reading: making a table from C cannot affect whether it is possible that a particular table be made from A. (The limit possibility claim.)
- The strong reading: making a table from C cannot affect whether a particular table is made from A; so any tables which can be made from A and C can be made together (in the same possible world). (The compossibility claim.)

The argument is only valid on the last, strongest, reading of the premise, but this seems too strong to me. The compossibility claim is defended by Guy Rohbraugh and Louis deRosset.[62] They argue for it via a principle they call the Inclusion Principle (IP) below [ibid. p711].

IP: Processes that turn hunks into tables seem to enjoy a form of independence from one another. A process that turns one hunk into a table need not interfere with any other.

Their idea here is two-fold. Firstly, the thought is that if a table T1 can be made from a hunk of wood H1 then the only way to prevent that possibility is to interfere with the process that leads to the production of T1 from H1. As they say, "each factor which prevents T1 from eventuating makes a difference to H1 or the people and tools involved in the productive effort."[63]

[62] [130]
[63] [ibid. p707]

Secondly, the claim, made explicit in IP, is that making a table from a hunk other than H1 need not so interfere. In that case, they argue, any two possible processes of making a table from a hunk may occur in the same world, if the hunks involved in the two process are wholly distinct; and while the processes *need* not result in the tables they actually result in (for Rohbraugh and deRosset do not want to assume Salmon's claim of the sufficiency of origin), they *might* result in those very tables. Now suppose, for reductio, that T1 could have been made from H2 ($\neq$H1). The process that turns H1 into T1 could run in the same world as the process that turns H2 into T2; and while the processes *need* not result in the same table they might do, since neither process need interfere with the other. Thus there is at least one possible case in which both processes are run and both result in T1. But this is impossible since one and the same table cannot be made from two distinct hunks of wood (and H1 and H2 are distinct in the world in question since they are actually distinct).

I do not find these considerations compelling; here is why.[64] We must, as before, distinguish between preventing the possibility of producing T1 from H1 via a process P in a world, and excluding the actuality of that situation. Producing T1 from H2 via P *excludes* the actuality of producing T1 from H1 via P in the sense that any world in which T1 is made from H2 via P is not a world in which T1 is made from H1 via P; but producing T1 from H2 via P does not *prevent* the possibility of producing T1 from H1 via P since in any world in which T1 is produced from H2 via P it is still *possible* in that world to produce T1 from H1 via P: possible, but not actual. I agree with Rohbraugh and deRosset's IP; processes turning hunks into tables need not interfere with each other. I also agree with them that the only way to *prevent* T1 originating from H1 via P would be to interfere with the process by which H1 comes to be made into a table — to interfere with P. But from this it follows only that in any world in which H1 is made into a table via P it is *possible in that world* that H1 be made into T1 via P, and this possibility is compatible with T1 in fact being made from some other hunk of wood. Rohbraugh and deRosset need a stronger principle for their argument to go through: that in order to *exclude* table T1 being made from H1 via P one needs to interfere with the process by which H1 is made into a table; i.e. the only way to prevent the actuality of T1 originating from H1 via P is to interfere with the process by which H1 is made into a table. But this principle is too strong, even by their

[64] I am indebted here to Sonia Roca, whom I worked with on a joint paper on this topic.

own lights. Remember that Rohbraugh and deRosset want to leave open the possibility that H1 be made into a table that is *not* T1 via P. Their argument, after all, is meant to be different from those like Salmon's which rely on the sufficiency of origin. But a world in which H1 is made into a table T2 via P ($\neq$ T1) is not a world in which H1 is made into T1 via P, since distinct tables do not originate from the same hunk via the same process. It is still *possible* in a world in which T2 is made from H1 via P that T1 be made from H1 via P, but making T2 from H1 via P excludes the actuality of making T1 from H1 via P; and yet there has been no interference with the process that turns H1 into a table. So Rohbraugh and deRosset are committed to the possibility that the making of T1 from H1 via P be excluded in virtue of T2 being made from H1 via P; why then should they find anything worrying about the making of T1 from H1 via P being excluded in virtue of T1 being made from H2 via P? If making a different table from the same hunk and process can rule out the making of the original table from that hunk and process, why is there anything problematic in the making of the same table from a different hunk and process giving rise to the same consequence? I think Rohbraugh and deRosset are making a similar mistake to Kripke. What would be problematic is if the *possibility* of making T1 from H1 via P was somehow ruled out by the making of T1 from H2 via P, or indeed the making of T2 from H1 via P. But the possibility of making T1 from H1 via P is not ruled out by either situation; only the actuality of it, which is ruled out by both.

Chapter 4

Truthmakers

> "When a statement is true, there is, *of course*, a state of affairs which makes it true."
> — J.L. Austin.[1]

I have argued that there is no particular problem in accounting for the source of true essentialist claims; we get essentialist claims for free given the underlying modal facts concerning what qualitative situations could have obtained. But what grounds these underlying modal facts?

Truthmaker theory offers a general approach to identifying the source of truth. According to truthmaker theory truth has its source in being: the idea is that when there is truth, there is some thing (or are some things) in virtue of which the truth is true. In this chapter I investigate truthmaker theory as means of identifying the source of modal truth. More precisely, the goal in this chapter will be to find out whether we might locate truthmakers for modal truths in the ontology of the actual, leaving the question as to whether or not there are non-actual truthmakers until the next chapter.

There are some who think that the project is not one that needs to be engaged in. Hugh Mellor for example, while a truthmaker theorist, thinks that there is no call for modal truthmakers. He says[2]

[1] [11, p23]
[2] [90, p213]

> [B]ecause the identity of a necessary proposition entails its truth, I cannot see why any other entity must exist to make it true. So, in particular, since any contingent proposition 'P' is *necessarily* contingent, I . . . see no need of a truthmaker for the necessary truth that P is contingent and hence that, in this sense, ¬P is possible.

Mellor's claim is that 'what grounds the truth of P?' is not a good one when P is necessary. The thought behind that, I take it, is that in demanding a grounding for the truth of P we are demanding a reason why the world is one way rather than another; if the world simply *had* to be that way then there is nothing that needs to be explained. For example, it has often seemed an important question to philosophers why there is something rather than nothing; but the problem has also been thought to lose its bite if we think that there *had* to be something.[3] I have my doubts, however, that the demand for explanation in general vanishes when the proposition to be explained is necessary; it seems that we can provide an illuminating explanation for why the square root of 2 is irrational (namely a proof that it cannot be written as m/n) even though that fact is necessary. And if it is not in general true that the demand for explanation is only appropriate for contingencies then I see no reason to think that the demand for truthmakers is only appropriate for contingent truths.

But even granting that truthmakers are not needed for necessary truths, we only avoid the demand for truthmakers for modal truths if S5 is true. The truths whose source I am seeking are not truths of the form P, where P happens to be necessary, but truths of the form 'Necessarily, P', or 'Possibly, P'. Truths of the latter form are only necessary according to the S5 system of modal logic. Mellor's reasons for rejecting the demand for modal truthmakers are because he is an S5-ist; he believes propositions of the form 'Necessarily, P' or 'Possibly, P' are themselves necessary, and so do not need to be grounded in being. I have doubts as to the truth of S5, and at any rate I don't think that we should *presuppose* S5 when attempting to locate the source of necessity, and so I reject Mellor's claim that no grounding is needed for modal truths (at least, I reject it when made for *that* reason).

I think it is worthwhile, then, to look for truthmakers for modal truths. Before I attempt this, however, I want to investigate the theory of truthmakers in general, for there is little point in

[3] See [82].

pinning our hopes on truthmaker theory if the theory is no good. I will start then by arguing for a rather unorthodox truthmaker theory which denies two traditional principles of truthmaker theory [Section 4.1]. I will then defend truthmaker theory against what I think is the biggest objection to it: that it brings an un-Humean commitment to necessary connections between distinct existences [4.2], and once I have done this I will investigate the prospect of truthmakers for modal truths [4.3].

Let us introduce a way of referring to truthmakers without prejudging, at this stage, what truthmakers are. Let the term 'TM(p)' refer (rigidly) to the truthmaker for some truth p if there is one. 'p' ranges over propositions, which I take to be the sole bearers of truth and thus the only things which are made true by truthmakers. I do not wish to make a stand on a particular view as to what propositions *are*; they may, for my purposes here, be sets of possible worlds, or they may be primitively representational entities. All that matters to me is that propositions are what we express with our utterances when we express something which is truth-apt.

The paradigmatic truthmaking relation is that which holds between a thing and the proposition that it exists.[4] Let us refer to the proposition that is expressed by a sentence S as [S], then we may write this constraint as $\forall x(\mathrm{TM}([\mathrm{x\ exists}]) = \mathrm{x})$

A note of caution before we go any further. I am speaking of *the* truthmaker for a proposition. This is misleading in two ways. First, it suggests that there is (at most) only one truthmaker for a true proposition. But I think that is false; the proposition that there are humans has many truthmakers — every human makes this proposition true. Also, it suggests that the truthmaking relation always holds between one thing and a proposition; but I want to leave it open that a proposition can be made true by a plurality of things. For example, the proposition that there are many galaxies in the universe is not made true by any one galaxy, and to say that it is made true by a mereological sum of some galaxies is to prejudge issues about composition; I prefer to say that the proposition is made true by a *plurality* of galaxies. Nevertheless, I will continue, for simplicity, to speak of *the* truthmaker for a proposition; but this should not be taken too literally.

[4] See [8, p12]

Why talk about truthmakers? The main historical motivation is due to C.B. Martin. Martin offered what seems to many to be a damning objection to Rylean behaviourism: namely that there is something wrong with attributing brute dispositional powers to a thing. What is wrong is that that leaves us true counterfactuals without anything in the world grounding their truth.

Ryle needs to account for the fact that we can have beliefs, desires, etc that are never manifested in our behaviour, and he does this by appealing to our dispositions to behave in a certain way. So although Joe's belief that Abraham Lincoln never ate a doughnut on the moon is never actually manifested in his behaviour, that he has such a belief is true in virtue of the fact that he is disposed to answer 'no' when faced with the question 'Did Abraham Lincoln ever eat a doughnut on the moon?' In that case the following counterfactual is true: if Joe were asked 'Did Abraham Lincoln ever eat a doughnut on the moon?' he would answer 'no'. But what, Martin asked, grounds this counterfactual? There is nothing in the world that accounts for its truth; there is no mental state that accounts for its truth because, according to Ryle, there are no mental states; but nothing except a mental state could ground a truth about Joe's beliefs. The counterfactual must be brute then; it is not true in virtue of anything which exists in the world, it is simply true. And the dissatisfaction with this appeal to a brute truth has led many to the truthmaker principle. What is wrong with Ryle's account is that there is no truthmaker for the counterfactual.

As Armstrong puts it[5]

> Ryle . . . leaves his counterfactuals hanging in the air. . . [H]e seems to think that he does not have to say what it is in the world that makes these counterfactuals true. . . What is the ground in reality, . . . the *truthmaker* for their truth?

Ted Sider advances the truthmaker argument against presentism. He says[6]

> The presentist claims that 'WAS(there exist some dinosaurs)' is true. But if there do not exist any past dinosaurs, what *grounds* the truth of this sentence? . . [T]he point [behind the truthmaker principle] is to rule out dubious ontologies that posit

[5] [7, p64]
[6] [135, p35-7]

> 'ungrounded' truths, for example 'brute counterfactuals' with no basis in the way things actually are. The thought is that it is illegitimate to postulate truths that 'float free' of the world. At first glance it would appear that the presentist's tensed truths float free of the world — they seem not to have truth-makers . . . For the presentist, all states of affairs are *currently* existing states of affairs, and the properties and relations of objects are confined to those of *currently* existing objects. But surely the truth about the past is not fixed by such facts about the present.

And a truthmaker argument has also been made against phenomenalism, the doctrine that physical objects are mere constructions out of sense-data. Against phenomenalism Armstrong says (again citing Martin as his inspiration)[7]

> Phenomenalists had a problem about physical objects and events at times that they are not being perceived. The solution . . . involved an appeal to certain *counterfactual* truths . . . of what perceptions would have been had, contrary to fact, a suitable observer had actually perceived them. . . Martin asked a simple question that seemed to go to the heart of the problem. Suppose that the required counterfactual propositions were indeed true. What are the truthmakers for these truths? Must there not be *some way the world is* in virtue of which these truths are true? *What is it?* How does the world make these truths true?

If you have any sympathy with the claims here that there is something objectionable about Rylean behaviourism, presentism or phenomenalism, as I do, then truthmaker theory starts to look quite attractive.

4.1 Truthmaker Necessitarianism and Maximalism

I want to consider now the principle of truthmaker necessitarianism: that the existence of a truthmaker necessitates the truth of the proposition it makes true. In symbols

$\Box((\text{TM}(p) \text{ exists}) \rightarrow p)$

[7] [9, p1]

It is sometimes suggested that the relationship between a truthmaker and that which it makes true is entailment. John Fox for example says "[B]y a truthmaker for A, I mean something *whose very existence* entails A."[8] But this can't be right. Entailment is a relation that holds, or fails to hold, between propositions. Entailment is necessary truth-preservation; p entails q if and only if p cannot be true and q false. Since only propositions are truth-apt, this only makes sense if p and q are propositions. Thus it makes no sense to speak of entailment holding or failing to hold between that which is not a proposition and something else; hence, since in general truthmakers are not propositions, the relation between a truthmaker and that which it makes true is not entailment. Propositions are *sometimes* truthmakers of course, if they exist at all; by the paradigmatic truthmaking relation the truthmaker for a claim that the proposition p exists is just the proposition p. But in general the truth-bearers are not the truthmakers.

The relationship between a truthmaker and that which it makes true is not entailment because it holds between the wrong kind of things. It is a cross-categorical link between something in the world and a proposition.[9] However, although the truthmaker does not entail the truth of the proposition it makes true, if truthmaker necessitarianism is true then the proposition that the truthmaker exists does.

I.e. [TM(p) exists] $\models$ p

That that follows from truthmaker necessitarianism is easily seen. Suppose that it is necessary that if the truthmaker for p exists then p is true. It is necessary that if the proposition that the truthmaker for p exists is true then the truthmaker for p exists; hence (by the transitivity of strict implication) it is necessary that if the proposition that the truthmaker for p exists is true then p is true, which is just to say that the proposition that the truthmaker for p exists entails p.

Truthmaker necessitarianism is standard in the truthmaker literature, although it has its detractors.[10] I am going to join the detractors. I will argue that acceptance of truthmaker necessitarianism is bound up with acceptance of truthmaker maximalism, the doctrine that every truth has a truthmaker. I will argue below that this latter doctrine should be abandoned,

[8] [43, p189]
[9] See [8, p12]
[10] See [90, p213-214], [54, p62-65] and [53], [97].

and with it truthmaker necessitarianism. But before I do, let us look at the consequences of retaining these two doctrines.

4.1.1 Negative and 'that's all' facts

Consider the truth that every cordate is a renate. What makes that true? If truthmaker necessitarianism is true then it cannot simply be the sum of each of the truthmakers for the claim of each particular cordate that it is a renate, for those things could exist even if there were something else which is a cordate but not a renate. The problem is that in general when there is a universal generalisation 'all Fs are G', it could be that every actual F is G and yet there be some other thing, not one of the actual Fs, which is F but not G. What then can play the role of the truthmaker for universal generalisations? How, if truthmaker necessitarianism is true, are we to account for the truth of [every cordate is a renate]? There seem to be two options.

1. We admit negative entities into our ontology: The truthmaker for [every cordate is a renate] is the negative fact that there is no cordate which is not a renate, or the state of affairs that there are no cordates which are not renates, etc.

2. We admit higher-order entities: [Every cordate is a renate] is made true by the truthmaker for each particular cordate that it is a renate and the second-order fact that *that's all the cordates there are*, or the state of affairs that has as its component the state of affairs of each actual cordate being a renate, and says of that state of affairs that it leaves no cordate out.

I am discounting a possible third option, and that is to claim that the world itself is the truthmaker for universal generalisations. This is not an option for someone who thinks that every maximal de dicto possibility is a de re way the actual world might be, since it is a de dicto possibility that there is an F which is not a G, for some true universal generalisation 'all Fs are G'. But it is open to one to deny that the de re modal properties of the world match the de dicto possibilities in that way: it is open to one to think that the world is essentially as it is; that is to say, every proposition that is true of the world is de re necessary of the world. In counterpart theoretic terms, the only counterpart of the actual world is itself. This does not entail that every

truth is necessary. There are still other worlds at which some actual falsehoods are true; but they are not counterparts of the actual world. In that case the world is a satisfactory truthmaker for 'all Fs are G', since the world cannot exist and that be false. Indeed, the world would be a satisfactory truthmaker (but not, in most cases, a minimal truthmaker) for any truth, since on this theory the world is essentially such that P, for all true propositions P. But that would be rather gratuitous essentialism; so while I think it is a coherent theory I will assume its falsity for our purposes here.

It is the latter of the above two options, the appeal to higher-order ontology, that Armstrong, a defender of truthmaker necessitarianism, opts for; his reason being that he thinks one needs higher-order states of affairs even if one has negative states of affairs. His argument does not, I think, depend on his belief that truthmakers are states of affairs and will, if sound, generalise. The argument goes as follows.[11]

Consider a world W in which there are two individuals a and b, such that a is F and b is G. F and G may be thought of here either as universals which a and b instantiate respectively, or as tropes which a and b have respectively; the decision will not matter for the purposes of this argument. In addition to the two truths in W [a is F] and [b is G] there are the two further truths [a is not G] and [b is not F]. What are the truthmakers for these two latter truths? No sum of the truthmakers for the former two if truthmaker necessitarianism is true. No matter whether the truthmaker for [a is F] is the state of affairs of a being F, or the mereological sum of a and the trope of a's F-ness or whatever, the existence of that thing is compatible with a being G. As before, we have two options then; we can opt for negative or higher-order truthmakers. On the former strategy the truthmaker for [a is not G] is the state of affairs of a not being G (or the mereological sum of a and the trope of a's non-G-ness, or whatever), on the latter strategy the truthmaker for [a is not G] and [b is not F] will be higher-order state of affairs that says that the first-order states of affairs that a is F and that b is G are all the first-order states of affairs there are (or a higher-order trope — the property of being all the first-order tropes there are — that belongs to the sum of a's F-ness and b's G-ness, or whatever). That second-order state of affairs necessitates the truths [a is not G] and [b is not F] because in every world in which it

[11] C.f. [6, p135]

exists neither the state of affairs that a is G nor the state of affairs that b is F exists; but those states of affairs exist in every world in which a is G or that b is F respectively. Now suppose we opt for the former strategy and make appeal to negative ontology to ground the negative facts [a is not G] and [b is not F]; still we must make appeal to higher-order ontology to account for the truth that those four states of affairs *are all the first-order things there are.*

The thrust of the argument is basically that we need higher-order ontology to account for truths of the form [there are only n first-order things]. It is no good to add a negative thing to our ontology which necessitates that there are no first-order things which are not among the n first-order things listed, for that negative thing would itself be a first-order thing and we would simply have shifted the problem to account for the truth [there are only n+1 first-order things]. To put the point perhaps more clearly: a negative fact could suffice for the truth of [all Fs are G] because it is not itself one of the Fs. The negative fact has as its constituents all the positive facts of the form 'this F is G', and says that there are no other such facts. But a negative fact cannot suffice for the truth of something of the form 'these are all the first-order things there are' since it itself is a first-order thing. A state of affairs cannot have itself as a constituent, so there cannot be a negative fact that says of itself and all the other first-order things that there are no others.

So there needs to be some higher-order thing to account for the truth that all the first-order things there are are all the first-order things there are. Since we need higher-order entities no matter what, we might as well get rid of negative entities and let the higher-order entities do the work of accounting for the truths [a is not G] and [b is not F].

There appears to be a risk of regress here: once we admit second-order things we look to be on the first step down the road of admitting third-order things, fourth-order things, and so on ad infinitum. Because there needs to be a truthmaker for a truth of the form 'such and such are all the second-order things there are', 'such and such are all the third-order things there are' etc. That regress wouldn't be vicious in the sense that it leads to the theory being incoherent; but it is certainly a vice on the grounds of (both qualitative and quantitative) parsimony.

But Armstrong does not think such a regress is opened up. He says of the world described above that "the third-order truth that these are all the first- and second-order truths [supervenes

on the first- and second-order truths], thus dealing with the apparent regress of higher-order truths."[12] So Armstrong thinks that given the first- and second-order things in a world, those things couldn't exist in a world in which some actual truth was false (or some actual falsehood true). Is he right? Call the set of all the first- and second-order things in a world w, S. Certainly in all the worlds in which the members of S exist there can be no difference in the first-order facts between that world and w. For if there was a difference in first-order facts then either one of the first-order states of affairs in S wouldn't exist, in which case not all the things in S exist, or some other first-order state of affairs, one that is not in S, would exist, in which case the second-order state of affairs in S wouldn't exist since, necessarily, it exists only if all the first-order states of affairs in S are all the first-order states of affairs there are. Either way then, something in S wouldn't exist, contrary to our initial assumption. So all the first-order facts must be same when all of S exist. What of the second-order facts? Could all the members of S exist in a world v and there be a difference in second-order facts between v and w? No; because all the second-order facts are facts concerning the first-order things, and they are all entailed by the actual first-order things being all and only the first-order things. First-order states of affairs are not the type of thing that instantiate properties; they are not red, or round, and they have no smell. The only facts concerning them are what ones exist, and what is entailed by that, such as that they are constituted a certain way. (The constitution of a state of affairs is essential to it; so that the state of affairs of As being red exists entails that there is a state of affairs constituted from A and redness.)

There is only ever one second-order thing then; and it suffices for all the second-order facts. (Indeed, it suffices for all the first-order facts since its existence necessitates the existence of the first-order things which are the truthmakers for the first-order facts.) Since there is only ever one second-order thing we do not need a truthmaker for the fact that such and such a second-order state of affairs is the only second-order thing there is; that second-order thing couldn't exist and that be false. So there only ever need be one second-order state of affairs to suffice for the truth of all second-order facts; and there cannot be a difference in what second-order facts are true and this second-order thing still exist. So Armstrong is right that there is no regress; we need go no higher than second-level ontology.

[12][ibid. p134]

Still, higher-order entities are too much for some to stomach. Armstrong admits that their admission into our ontology is "objectionable . . . a major sin against economy."[13] Armstrong accepts higher-order things because without them the principles of truthmaker theory do not come out true. What options remain for those who find the postulation of a higher-order ontology too objectionable to make their benefit in truthmaking worth it?

4.1.2 Abandoning Maximalism and Necessitarianism

How did we get into the problem of requiring higher-order ontology? It was because we were seeking a truthmaker for a truth of the form 'such and such are all the (first-order) things there are'. But that argument, of course, presupposes that there needs to be a truthmaker for a truth of that form. Armstrong requires there to be a truthmaker for such truths because he is a truthmaker maximalist; he believes that *every* truth has a truthmaker. That is what I want to deny. Armstrong says that to abandon truthmaker maximalism would be "to abandon ontological seriousness."[14] His worry, I take it, is that once we admit that some truths don't require truthmakers then we are in no better a position than Ryle was with his brute counterfactuals. That worry is a serious one, but not, I think, a damning one. It poses a challenge to the theorist who wants to deny truthmaker maximalism: they must give us non ad-hoc reasons for why certain truths don't require truthmakers if they are to avoid the objection that they have abandoned ontological seriousness.

Why might some truths not require truthmakers? A natural thought is that a proposition does not require a truthmaker if it is a truth-functional construction out of atomic propositions. This is the position advocated by Hugh Mellor.[15] He says

> Some . . . truths need no truthmakers, notably true truth-functions, whose truth follows from the truth values of their constituents. We may *say* of course that 'P&Q' and 'P∨Q' are 'made true' by the truth of 'P' and 'Q'; but this is just the entailment of one proposition by others, not the 'cross-categorical' link between propositions and other entities that concerns us here. That is what true truth-functions do not need

[13] [ibid.]
[14] [6, p135]
[15] [90, p213]

> and therefore, I claim, do not have. . . In particular, negative propositions do not need them, since if 'P' is made true by S, all it takes to make 'P' false and hence '¬P' true is that S *not* exist.

I think Mellor is right about this. There is no need to look for the truthmaker for a true conjunctive proposition; it is true in virtue of there being a truthmaker for each conjunct. If there is a truthmaker for Q then there is no need to look for a truthmaker for P → Q; the truthmaker for Q suffices. And, most importantly, there is no need to look for the truthmakers for negative truths — a negative proposition is true just in case the corresponding positive proposition lacks a truthmaker.[16,17]

This is not ad-hoc, and so I think it avoids Armstrong's worry. We are not abandoning the search for truthmakers for these truths because it has become too hard to find them; it is perfectly reasonable to hold that true truth-functional propositions require no truthmakers; all that it takes to fix the truth-value of truth-functional constructions of atomic propositions is to fix the truth-value of its constituents.

Once we thus abandon truthmaker maximalism there is no need to make appeal to negative or higher-order ontology. The truth in W above that a is not G is not made true by the state of affairs of a's not being G nor is it made true by the higher-order state of affairs that says that the states of affairs of a's being F and of b's being G are all the first-order states of affairs there are; it is not made true by anything. It is true not in virtue of the existence of anything but in virtue of the non-existence of a truthmaker for [a is G].

Giving up truthmaker maximalism results in our abandoning truthmaker necessitarianism.[18] Consider a world W containing two individuals a and b, which are both F, and nothing else other than what is entailed by this. What makes it true in W that everything is F? Only this: the truthmakers for [a is F] and for [b is F]. Those truthmakers do not *necessitate* the truth that everything is F of course; whatever they are, they could coexist with some thing which is not F,

[16] It is important to note, as Mellor points out [loc cit.], that no claim is being made to the effect that we can always tell which, if either, of 'P' or '¬P' are negative propositions.

[17] This position on negative propositions rules out a treatment of indeterminacy whereby the indeterminate propositions are said to be precisely the ones which both lack a truthmaker and lack a truthmaker for their negation. I am happy to rule this position out.

[18] C.f. Mellor [op cit. p214]

but the fact that there is nothing in W which is not F is a negative fact which does not itself require a truthmaker. a and b are the only things in W[19], hence [everything is F] is true solely in virtue of that which makes a F and that which makes b F; there is no need to make appeal to anything else. If God were to make W actual He would make a, make b, make them both F, and then stop; that would be all He had to do to make it true that everything is F; He would not need to add to W a negative or higher-order thing.

Consider another example. In the 2004 annual Logic and Metaphysics versus Moral Philosophy football match in St Andrews, MP beat L&M by 5 goals to 4. The truthmaker for [MP won] is just the truthmaker for [9 goals were scored], since *all that had to happen* for MP to win the match is that the goals which were actually scored were scored; MP's victory was not a consequence of their learning how to create the 'and that's all the goals that were scored' entity; MP's victory was brought about just by the respective teams scoring the goals they did. The truthmaker for [9 goals were scored] is the truthmaker for [MP won], but while the existence of that truthmaker necessitates the truth of the former proposition it fails to necessitate the truth of the latter, since it could exist in a world in which L&M scored 2 more goals than they actually did. So truthmaker necessitarianism has to go; but this departure should not, I think, worry us. We can still hold on to a restricted form of truthmaker necessitarianism; for it is still true that the proposition that A exists entails the truth of p where p is an *atomic* proposition and A its truthmaker. We are only abandoning the claim that the existence of TM(p) necessitates the truth of p when p is a non-atomic proposition; and the failure of necessitation is not through any insufficiency of the truthmaker, but due to how the truth of non-atomic propositions depends on its component parts.

Some truthmaker theorists abandon truthmaker necessitarianism even for atomic propositions. Josh Parsons[20] for example says that the truthmaker for p is that which is intrinsically such that p. So the truthmaker for (the atomic proposition) [A is red] is just A. But A does not have its intrinsic properties essentially, so the existence of A fails to necessitate the truth of that which it makes true. Parsons likes this theory because it is a theory of truthmakers which is satisfactory

[19] Well, there are the truthmakers as well, of course, and any necessary existents, but suppose our quantifier is appropriately restricted.

[20] [97]

to the nominalist (or better: what Armstrong calls the extreme nominalist; not just someone who doesn't believe in universals but who doesn't believe in properties at all). The truthmakers, for Parsons, are ordinary individuals. The reason I don't like this, the reason I want to insist that truthmakers for atomic propositions necessitate the truth of that which they make true, is partly methodological: they concern my views on how we should proceed in identifying truthmakers. It seems to me that we should look to what *kinds* of thing perform the truthmaking role in the simple cases and then see whether or not there are things of the same kind that can perform the role in the harder cases (such as, e.g., modal facts).

Now our problem, remember, is to find truthmakers for necessary truths. And so we should proceed thus: identify the truthmakers for atomic propositions and see whether there is anything of a kind to those things that are adequate truthmakers for modal truths. Now we will need some guiding principles to work out what the truthmakers for atomic propositions are, and one of these, I think, is necessitarianism. Because we need some way to get a grip on the truthmaker relation; we need some way of making sense of the 'true in virtue of' relation. And we get a grip on this relation only by considering the relation of necessitation between a thing and the truth of a proposition. Now Armstrong takes this as evidence that in all cases the existence of a truthmaker necessitates the truth of that which it makes true. But I think that is unnecessary. Consideration of the relation of necessitation lets us identify the truthmakers for atomic propositions; and then, given this new data (i.e. what kinds of thing truthmakers are), we can figure out, hopefully, what the truthmakers are in the more difficult case of non-atomic propositions. And if necessitation fails in these more difficult cases then so be it. We have not lost our grip on the truthmaking relation because now we have evidence that the truthmaking relation is that which holds between things of a certain kind, perhaps states of affairs, and propositions. This is why, although I abandon truthmaker necessitarianism in general, I hold it with respect to atomic propositions, and why I cannot accept Parsons' theory. Parsons' has given us no reason to think that the relation that holds between that which is intrinsically such that p, and p, is the truthmaking relation; so we have no reason to think that Parsons is not simply using the term 'truthmaker' in a way different to us. But that the relation of necessitation holds between things of a kind K and atomic facts is good evidence that the relation between

the things of kind K and facts is the truthmaker relation. That is to say that the truthmaker relation is that relation which holds in general between things of kind K and facts, atomic or otherwise; and there is no problem if the relation sometimes fails to be one of necessitation.

I can offer a pleasing story about *why* necessitarianism fails. When the existence of TM[p] fails to necessitate p that is because the truth of p supervenes not just on TM[p]'s existence but also on the falsity of certain negative propositions: propositions such that were they true, TM[p] could exist and p would be false. Call the negations of these negative propositions *releasing factors* for p.

The truth of p supervenes on the existence of TM[p] and the falsity of the releasing factors for p. But that the releasing factors are false is a negative fact which does not need a truthmaker. Call one such releasing factor for p R. TM[p] can exist in a world in which R is true, and in such a world p is false. But this is explainable; TM[p] does not necessitate the truth of ¬R. There exists no thing that necessitates the truth of ¬R. Nor need we demand such a thing, because ¬R is a negative proposition — it requires no truthmaker.

There are other intuitive examples of non-atomic propositions being made true by things that fail to necessitate their truth. Counterfactuals are among them. Here is an example from John Heil[21]

> Take the assertion
> (P) If you drank this cyanide-laced tea, you would die.
> Suppose (P) is true in virtue of some object or fact, a . . . Could we imagine a world that included a, but in which (P) was false?
>
> Think of a world that included the cyanide-laced cup of tea but included, in addition, your having in hand an antidote. In that case, (P) could be false despite the presence of a. . . . More generally, an assertion, A, might fail to hold, not because [the truthmaker for A] is absent, but because [the truthmaker for A] is accompanied by a defeater.

It should come as no surprise, given the above, that the truthmaker for a counterfactual should

[21] [53, p232-233]

fail to necessitate the truth of that which it makes true. Counterfactuals are a kind of universal generalisation: in *every* selected world in which p is the case, q is the case. (P) amounts to the statement that in every selected world in which you drink the tea, you die. It is true because of the absence of a selected world in which you drink the tea and live. But that there is no such selected world is a negative fact and not something that requires a truthmaker; it is sufficient that there is no truthmaker for the fact that there is a selected world at which you drink the tea and live.

Or if talk of worlds makes you queasy we can put the same point differently. The truth of (P) supervenes not only on facts concerning your disposition to die if you ingest cyanide, but also on the falsity of any propositions which would entail the might counterfactual: if you drink the tea you might live. R_1: [You have to hand a cyanide antidote] is one such proposition, since the presence of the antidote means that if you drink the tea you might then take the antidote, and hence might live. So the truth of (P) supervenes on both your dispositions and the falsity of R_1. And also, of course, on the falsity of R_2: [There is close by a very skilled doctor], and R_3: [There is a God who will miraculously save you from dying of cyanide poisoning] etc. But $\neg R_1$, $\neg R_2$, $\neg R_3$ etc are all negative facts. They have no truthmaker. A fortiori, a, the truthmaker for (P), is not their truthmaker, and hence there is no problem at all that a can exist in worlds in which $\neg R_1$, $\neg R_2$, $\neg R_3$ etc are false, i.e. worlds in which R_1, R_2, R_3 etc are true, i.e. worlds in which (P) is false.

So consider a world in which you have the antidote in hand. In that world there exists a truthmaker for the proposition that there is a selected world at which you drink the tea and live, or if you prefer: a truthmaker for the might counterfactual 'if you drink the tea you might live'. It is therefore false that if you drink it you will die. So (P) is false; nevertheless, the truthmaker for (P), a, exists at this world: if you drink the tea in this world it is just as true as in the actual world that you will be disposed to die. But since the existence of a only suffices for the truth of (P) given that there is no truthmaker for the fact that there is a selected world at which you drink the tea and live, no truthmaker for the might counterfactual, the existence of a does not suffice for the truth of (P) in this world.

In general a counterfactual $p \Box\!\!\rightarrow q$, unless the corresponding strict conditional is also true, is

only true in the absence of defeaters (releasing factors). There will be some contingently false r such that $\neg((p \wedge r)\Box\!\!\rightarrow q)$. But that each such r is false is not something which requires a truthmaker, so there need be nothing whose existence necessitates that r is false. A fortiori, the truthmaker for the counterfactual need not necessitate that r is false; hence its existence is compatible with a truthmaker for r. Hence its existence is compatible with the counterfactual being false. Likewise, just because the truth of the would counterfactual $p\Box\!\!\rightarrow q$ is incompatible with the truth of the might counterfactual $p\Diamond\!\!\rightarrow \neg q$, it does not follow that the existence of the truthmaker for the would counterfactual must necessitate that the might counterfactual is false. *Nothing* need necessitate that the might counterfactual is false, since that is a negative fact that does not require a truthmaker. The situation in general with regard to universal generalisations is this. What makes a universal generalisation true is just those things which make each of the instances true, if there are such things.[22] But the existence of the truthmaker for the universal generalisation is not sufficient for the truth of the universal generalisation; it only brings about the truth of the universal generalisation in worlds in which there is no truthmaker for the incompatible existential generalisation.

So the existence of the truthmaker for (P) is compatible with the existence of truthmakers for defeaters of (P), and thus with the falsehood of (P). Armstrong, of course, would use this as an argument that a cannot have been the truthmaker for (P). He says[23]

> If it is said that the truthmaker for a truth could have failed to make the truth true, then we will surely think that the alleged truthmaker was insufficient by itself and requires to be supplemented in some way. A contingently sufficient truthmaker will be true only *in circumstances that obtain in this world.* But then these circumstances, whatever they are, must be added to give the full truthmaker.

So for Armstrong the truthmaker for (P) must be a *and* the negative or second-order thing whose

[22] If each of the instances is a negative fact which requires no truthmaker then, on my view, the universal generalisation requires no truthmaker. This implies that negative existentials sometimes require truthmakers and sometimes do not. The negative existential $\neg\exists x(Fx)$ requires a truthmaker iff the corresponding universal generalisation $\forall x\neg(Fx)$ requires a truthmaker. So [there are no chimeras] does not require a truthmaker because [everything is not a chimera] does not require one, since no instance of 'a is not a chimera' requires a truthmaker. But [there is no cordate who is not a renate] does require a truthmaker; and it is made true by that (or those things) which make true for every cordate that it is a renate.

[23] [6, p116]

existence necessitates the non-existence of the antidote (and any other defeater).

Armstrong says that the extra circumstances which are required for the truth of p in addition to the existence of some thing A must be added to A to give the full truthmaker for p. But this will not be accepted by the denier of truthmaker maximalism, because they will resist the reification of such circumstances that this comment implies. So Armstrong's argument establishes only that *if one is a truthmaker maximalist* then one should be a truthmaker necessitarian. In Mellor's world W, TM([a is F]) and TM([b is F]) do not necessitate the truth of [everything is F]; TM([a is F]) and TM([b is F]) only lead to the truth of [everything is F] in the circumstances that there is nothing which is not F. But to demand that those circumstances must be added to TM([a is F]) and TM([b is F]) to give the full truthmaker for [everything is F] is to reify those circumstances in a way the opponent of truthmaker maximalism will not accept. The circumstances can only be added to the truthmaker *if they exist*, but to deny the existence of higher-order or negative ontology — as the denier of truthmaker maximalism does — is precisely to deny that there *are* any circumstances answering to the description 'there is nothing which is not F'. That's not to say that we can't truly describe a situation thus, of course, merely to deny that there is any ontology in the world which corresponds to the description. To make this denial is not to go against the correspondence theory of truth; merely to go against the view that what our true propositions correspond to can simply be read off from the sentences which we use to express them. Ontology is hard; it is not to be done simply by observing how we speak.

I conclude that in the case of non-atomic propositions there is no requirement for the truthmaker to necessitate the truth of the proposition in question, and hence that there is no need for the truthmaker theorist to make appeal to strange negative or higher-order ontology.

4.2 Necessary Connections

4.2.1 Lewis on Necessary Connections

> "[T]he demand for truthmakers just *is* a demand for necessary connections." — David Lewis[24]

[24] [72, p219]

The above quote aptly summarises the reason Lewis for so long rejected truthmaker theory; namely, that to accept the theory is to violate the principle that "gives us our best handle on the question [of] what possibilities there are"[25], namely the Humean denial of necessary connections between wholly distinct contingent[26] existents: the principle that "anything can coexist with anything else . . . Likewise, *anything can fail to coexist with anything else.*"[27]

Consider some contingent atomic proposition of the form [A is F]. This proposition, says the truthmaker theorist, has a truthmaker which necessitates its truth. The truthmaker for [A is F] is not A, since A might not have been F. Nor is the truthmaker for [A is F] the sum of A and any other thing C which on its own would not suffice for the truth of [A is F], since this sum could also exist in a world in which A is not F. The truthmaker for [A is F], then, must be something wholly distinct from A, call it B. But now we have a necessary connection between wholly distinct existents; for any world in which B exists is one in which A is F, and any world in which A is F is one in which A *exists*. So there is a necessary connection between the existence of B and the existence of A. Lewis finds this necessary connection mysterious, and rejects truthmaker theory because of it. As he puts it[28]

> B is entirely distinct from A . . . yet B's existence is necessarily connected to whether or not A has F. Necessarily, if A has F . . . then B must exist; necessarily, if A had lacked F . . . then B would not have existed.

Actually, we're under no pressure to accept the first of these necessary connections: "Necessarily, if A has F . . . then B must exist". Because we need not, and probably should not, accept the doctrine that there is only one possible truthmaker for any truth. A could be F and B not exist, since there could have been something else that made it true that A is F. But the second of the necessary connections is good: if A had lacked F then B would not have existed, because if B exists then it must be the case that A is F.

For an example, consider the view that truthmakers for propositions of the form [x is F] are tropes, particularised properties, that belong to substrata. The truthmaker for [A is F], on this

[25] [74, p611]
[26] I won't keep making this qualification. It is to be assumed throughout this section that we are dealing with contingent existents.
[27] [66, p87-88]
[28] [72, p215]

view, is the trope of A's F-ness. If this is right then the trope must be *non-transferable*; it must essentially attach to A in the sense that no individual which was not A could have it; for otherwise the truthmaker for [A is F] could exist in a world in which A was not F, which is impossible. In that case there is a necessary connection between the existence of A's F-ness and the wholly distinct thing which is A's substratum.

Or consider the view that the truthmaker for [A is F] is the state of affairs of A being F. In every world in which this state of affairs exists A is F, and so A exists. Is the state of affairs wholly distinct form A? Yes — at least in the sense Lewis means. It is true that the state of affairs has A as a *constituent*, but it does not have A as a *part*. The relationship between a constituent of a state of affairs and the state of affairs is not the part/whole relationship since all the constituents could exist — and hence the sum of all the constituents could exist[29] — in a world in which the state of affairs does not exist. So the state of affairs of A being F is mereologically wholly distinct from A and hence we have, in the intended sense, a necessary connection between wholly distinct existents.

Of course, it is open to someone to deny Lewis' reading of Hume's maxim, and to claim that the above necessary connection is harmless since, while the state of affairs is *mereologically* distinct from A, it involves A in a very intimate way. But let us grant Lewis' reading of the principle for now and see where it leads us. It seems that we are forced to make a choice between truthmaker theory or the Humean denial of necessary connections. Or are we?

4.2.2 Counterpart Theory to the rescue?

Lewis was later to abandon his rejection of truthmaker theory, because he thought his counterpart theory could let him hold both truthmaker theory *and* Hume's denial of necessary connections.[30] The problem arose, remember, because objects do not have their properties essentially, thus we cannot claim that in general a thing A is the truthmaker for [A is F] since A might not have been F. Of course, once one goes counterpart theoretic, an object is only essentially or not essentially such and such *relative to some counterpart relation*. It may well be true that under

[29] The argument here assumes the possibility of unrestricted composition.
[30] See [75] and [77]

the counterpart relations invoked when I speak of A in an everyday context that A might not have been F, but I can invoke a counterpart relation under which A is essentially F by referring to A as 'A-qua-F'. A-qua-F is essentially F, so its existence necessitates the truth of [A is F]. So we can have a truthmaker for [A is F] without committing to necessary connections between distinct existences, for A-qua-F is nothing but A itself; merely A considered under a counterpart relation according to which A is essentially F.

There is an initial worry with this approach that I will not dwell on. Lewis is presupposing that we are able to invoke a counterpart relation according to which A is essentially F in the same context in which we are trying to determine what the truthmaker for [A is F] is. This is far from obvious. Why should we not think that in invoking such a counterpart relation we have simply changed the context from a context in which we are talking about truthmakers. Isn't the fact that we don't think that A is an adequate truthmaker for [A is F] (and that Lewis thought so for so long!) a datum in evidence of the claim that in a context when we are talking about truthmaking A is not essentially F? If so there is no reason to think that we *can* invoke a counterpart relation according to which A is essentially F without changing the context. If this is right then the things qua truthmakers theory is mistaken. Ordinary objects cannot serve as the truthmakers for the fact that they are a certain way, for even if there is a counterpart relation according to which they are that way essentially, we cannot invoke such a relation without changing the subject. That is a worry; but put it aside for the sake of argument and grant Lewis the assumption that we can truly assert that A is essentially F in a context in which we are looking for the truthmaker for [A is F]. Even granting this his theory has problems.

Let us consider first a worry which Lewis considers and thinks he has an easy answer to. The problem is this: Lewis says that A-qua-Φ is the truthmaker for 'A is Φ', but surely B-qua-(A being Φ) would do just as well. But in that case any object can — when appropriately described — be the truthmaker for any truth whatsoever; Plato-qua-p will suffice as the truthmaker for any truth p. But that makes a mockery of truthmaker theory. And Lewis' easy answer is: the counterpart relation invoked by 'B-qua-(A being Φ)' or 'Plato-qua-p' is 'peculiar' in a way that the counterpart relation invoked by 'A-qua-Φ' is not.[31] What does this peculiarity amount to?

[31] [75, p32-3]

Lewis cites two ways in which the relation is peculiar. First, that the respect in which objects that are counterparts under these relations are similar is "one that would strike us in almost any context as an utterly unimportant similarity."[32] Second, that it is an "entirely extrinsic similarity".[33] I am dubious that anything of much philosophical importance can hang on the first consideration. What does it matter how the counterpart relations strike *us*? Whether an object is an adequate truthmaker for some truth p is surely independent of our reactions to hearing that object described in a certain way. If the counterpart relations invoked are to be deemed inadmissible on the charge of peculiarity then, I suggest, it is Lewis' second worry that must do the work. In that case one may justly ask: must an admissible counterpart relation judge *solely* on intrinsic similarity or merely take some intrinsic similarity into account? Lewis — in the postscript written with Gideon Rosen[34] — answers the latter in order to get truthmakers for negative existentials. The claim is that the truthmaker for 'there are no unicorns' is the world-qua-unaccompanied by unicorns; and although being unaccompanied is an extrinsic property, Rosen and Lewis are happy with similarity with respect to being unaccompanied marking out a genuine counterpart relation because "the property of being completely unaccompanied . . . does seem quite important to the character of anything that has it."[35] But the admission that the truthmaking counterpart relation can judge on matters of similarity which are in part a matter of external similarity makes it harder for Lewis to avoid the above problem that any object can be the truthmaker for any truth. Consider, for example, Socrates-qua-(snub-nosed and such that the Eiffel tower has 1665 steps); that invokes a counterpart relation that judges similarity on both intrinsic and extrinsic properties. In any world in which there is a counterpart of that object (under the invoked relation) it is true that the Eiffel tower has 1665 steps; but that object is just Socrates — and surely Socrates does not make it true, no matter how we describe him, that the Eiffel tower has 1665 steps.

I think we should take the following conclusion from this problem: that for the counterpart theorist, on the assumption that in the context of discussing truthmakers we can invoke a counterpart relation for any object x and any property Φ of x such that x is essentially Φ according

[32] [ibid. p32]
[33] [ibid.]
[34] [77]
[35] [ibid. p40]

to that relation, the truthmaking relation is simply not an interesting relation. The triviality of finding an object that, according to some counterpart relation, is essentially such that P is true, for any true proposition P, simply shows that the relation of an object necessitating some truth is not an interesting metaphysical relation. But not every counterpart theorist must give up on truthmaking as an interesting theory. I am inclined to think that in the context of determining what makes [A is F] true that A is not essentially F, and hence that A is not a suitable truthmaker for this atomic proposition. I think the triviality of finding a counterpart relation according to which A is essentially F just shows us that that relation is not what the truthmaker theorist was looking for and so is irrelevant in this context. Truthmaker theory was meant to cause problems for Rylean behaviourism and phenomenalism; but it doesn't if the goal is just to find some counterpart relation according to which some actual thing necessitates the truths in question. Similarly, there is supposed to be a truthmaker argument against nominalism, the view that there are no properties; but there isn't if the goal is just to find a counterpart relation according to which objects have their properties essentially. So this simply can't be the goal of the truthmaker theorist: such counterpart relations must simply be irrelevant in this context. So I consider myself in the same boat as the trans-world identity theorist in trying to find some other kind of thing — perhaps tropes or states of affairs — whose existence, in this very context, necessitates the truth of [A is F].

How might such a counterpart theorist, or a trans-world identity theorist, respond to Lewis' claim that while traditional truthmaker theory violates the Humean denial of necessary connections, his things-qua-truthmakers version does not? The trans-world identity theorist in particular might worry about whether we really *have* avoided necessary connections between distinct existences by adopting counterpart theory. Some might say that we have embraced them; after all, in counterpart theory every thing that exists in a world exists *only* in that world. There is no strict trans-world identity. In that case don't we have necessary connections galore? For any actual object x, x only exists in worlds in which all its actual worldmates exist, since x only exists in the actual world. The natural response to this thought, of course, is that one should interpret the denial of necessary connections counterpart theoretically. But to do so is no trivial matter, as we shall soon see.

4.2.3 Lonely and Independent existence

Consider again Lewis' statement of the Humean principle; he says "the principle is that anything can coexist with anything else . . . Likewise, *anything can fail to coexist with anything else.*"[36] The latter demand here is ambiguous: it could be read in two ways. It could be a demand that for any thing A, A can fail to coexist with any other *particular* thing, or it could be a demand that for any thing A, A can fail to coexist with anything *per se* (i.e. that A can fail to coexist with anything at all). That is, the demand could be understood as saying that there should be no thing B which is distinct from A and which exists in every world in which A exists, or the demand could be understood as saying that there are worlds in which A exists and in which no thing that is wholly distinct from A exists. The two competing principles can be formulated (in QML) as follows:[37]

$$HM : \forall x \neg \exists y (x \neq y \wedge \Box(\exists z(z = x) \rightarrow \exists v(v = y)))$$

$$DNC : \forall x \Diamond(\forall y(y < x))$$

I am more attracted to HM than DNC, since DNC rules out the existence of Aristotelian universals. On the Aristotelian conception of universals, universals only exist in worlds in which they are instantiated. Any world in which a universal exists, then, is a world in which it is instantiated. But universals do not instantiate themselves, particulars instantiate them. Particulars and universals do not overlap, so any world in which a universal exists is a world in which something wholly distinct from it exists. So we have a violation of DNC, but no violation of HM; there is no thing such that its existence is necessitated by the existence of the universal, since universals can be instantiated by more than one thing. That leaves it open that DNC is true when the quantifier ranges only over *particulars*; but even this is doubtful: it seems whenever there is a thing, there must also be the impure, and contingently existing, sets that have that thing in their transitive closure.[38] Thus nothing can exist unaccompanied: everything will be

[36] [66, p87-88]

[37] Where 'x<y' is read as 'x is a (perhaps non-proper) part of y', understood such that $\forall x(x < x)$ is true.

[38] Note that I am assuming only that in every world in which a thing A exists, there is a thing in that world that is the singleton of A in that world, and a thing which is the pair set of A and singleton A and . . . etc. I am *not* assuming that the thing which is *actually* the singleton of A exists in every world in which A exists. I am

accompanied by sets.

So I have doubts about DNC; but let us see how the counterpart theorist will view the relevant translation of DNC, and HM. Straightforwardly translating into counterpart theory would give us for HM that for any two things x and y there could exist a counterpart of x without there existing a counterpart of y, and for DNC that for all things x there could exist a world in which there is nothing wholly distinct from a counterpart of x. These are HM_{CT} and DNC_{CT} below.

$$HM_{CT} : a \neq b \rightarrow \exists w(\exists x(Ixw \wedge Cxa) \wedge \neg\exists y(Iyw \wedge Cyb))$$

$$DNC_{CT} : \forall x\exists w(\exists y(Iyw \wedge Cyx \wedge \forall z(Izw \rightarrow z < y)))$$

But neither of these can be true. DNC_{CT} cannot be true because there are counterpart relations which weigh heavily extrinsic similarity. As Lewis says "counterparts are united by similarity, but often the relevant similarity is mostly extrinsic."[39] For example, a counterpart relation is invoked when I consider counterfactual circumstances in which I was born at different times, in different places, perhaps from different parents etc. The important thing in each case is that I am considering myself as something which is *born*; thus it is natural to think that something will only count as my counterpart under this relation if it is something which was born. And hence it seems that in every world in which there is a counterpart of me under this relation there must also be objects which are the parents of that thing; for a thing cannot be born yet not have parents. But the parents of a thing cannot be a part of the child; at least, not so that the child can still be said to have been *born* in the sense we are interested in. Thus whenever a counterpart of me exists under this relation, something which is not a part of my counterpart must also exist.

HM_{CT} cannot be true because two distinct things can have the same counterpart, and they can be such that anything that will count as a counterpart of one under some counterpart relation will count as a counterpart of the other under that relation. Consider Max Black's homogeneous

leaving it open that the impure sets that have A in their transitive closure are not identical from world to world. This is relevant to the discussion of OP and OP* below.

[39][Lewis, op cit. p88]

iron spheres, Castor and Pollux, as an example.[40] Each sphere resembles the other precisely, down to the most minute detail. Therefore anything which resembles one sphere enough to be called its counterpart under some counterpart relation resembles the other sphere enough to be called its counterpart under that relation. So even though they are distinct, there cannot exist a counterpart of Castor without there existing a counterpart of Pollux. But this problem seems easily fixed; intuitively there is no problem if whenever there is a counterpart of Castor there is also a counterpart of Pollux, precisely because it is the one thing which is counting as a counterpart of them both. What would be objectionable to the Humean is not that there are two distinct things which must co-exist because anything which represents one as existing also represents the other as existing, but that there are two distinct things such that necessarily if something is a counterpart of one then something *else* is a counterpart of the other. So let us have another go at formulating HM in counterpart theory. What we want to rule out is that in every world in which there is a counterpart of a there is something wholly distinct from that counterpart which is a counterpart of b (such that a≠b); or equivalently, that there is a world which contains a counterpart of a and is such that any counterpart of b in that world is a (perhaps non-proper) part of the counterpart of a. That is HM^*_{CT} below.

$$HM^*_{CT} : a \neq b \rightarrow \exists w \exists x (Ixw \wedge Cxa \wedge \forall y (Cyb \rightarrow y < x))$$

Is HM^*_{CT} acceptable? I think not. As we saw above, set theory seems to falsify DNC because according to set theory objects cannot exist lonely — they must be accompanied by the pure sets that have them in their transitive closure. But set theory, together with some plausible essentialism, also seems to falsify HM^*_{CT}; for it seems necessary that whenever the singleton of a thing A exists, A exists as well. In counterpart theoretic terms, you can't have a counterpart of singleton A without having something wholly distinct from that counterpart which is a counterpart of A — namely, the member of that counterpart. This is precisely because a set is only a counterpart of singleton A if it is a singleton that has as its member a counterpart of A.[41] And perhaps this holds the other way as well: that necessarily whenever A exists, singleton A exists.

[40] [16]

[41] I am assuming that the members of a set are not mereological parts of that set.

(I say 'perhaps' for reasons that will become evident in the discussion of OP and OP* below.)

I should point out that the problems set theory brings for both HM and DNC, and their counterpart-theoretic translations, do not arise for a Lewisian with respect to worlds, since sets don't exist according to worlds in the way that concreta do: they are not part of the mereological sum that is the world. Since DNC only demands that objects can exist without worldmates, it is of no consequence to the Lewisian that necessarily whenever there is some object there are the impure sets with that object in their transitive closure, and similar remarks apply to HM. But I am not a Lewisian about worlds — worlds, I hold, represent sets as existing in the same way as they represent concreta as existing — and so I think there is a tension between set theory and DNC and HM.

4.2.4 Lonely and Independent Duplicates

HM and DNC, and their counterpart-theoretic translations, are unacceptable. Lewis, however, did not rely on any of them. Lewis instead appealed to the notion of duplication to do the job where the notion of a counterpart failed. The principle Lewis held is that for any object a there is a world at which there exists a duplicate of a and nothing else, where a duplicate of a is an object which is intrinsically identical to a. This is implied by Lewis' principle of recombination:

Com: For any wholly distinct things $x_1, x_2, ...x_n$ there is a world containing any positive number of duplicates of each, and no thing which does not overlap any of those duplicates, size and space permitting.[42]

There is a problem with how one could come to acquire justification for Com, which I will take a digression to discuss.

[42]The formulation is close to that given by Divers and Melia [35, p16]. One difference is that they do not add the 'and no thing which does not overlap any of those duplicates' qualification, but this is necessary if Com is to have the consequence, that Lewis takes it to have, that for any object there is a world containing a lonely duplicate of that object. The other difference I have made is to insist that the individuals $x_1, x_2, ...x_n$ are wholly distinct. The point here is to ensure that Com does not entail the existence of the impossible world where there are some positive number of duplicates of me and no duplicates of some proper part of me. This seems to have been Lewis' thought when he said "anything can coexist with anything else, *at least provided they occupy distinct spatiotemporal positions.*" (Lewis [op cit. p88], my emphasis.) I presume he means by distinct spatiotemporal positions non-overlapping spatiotemporal positions, in which case his qualification amounts to the demand for whole distinctness: if the spatiotemporal location of x does not overlap with the spatiotemporal location of y then there is no part of x that shares its location with any part of y; in which case x and y share no parts in common, in which case they are wholly distinct. For Lewis' original comments on recombination see [66, p86-92].

Digression: Recombination and Intrinsicality

> If any kind of combinatorial analysis of intrinsicness can work, we have to assume something like Hume's dictum that there are no necessary connections between distinct existences. . . This might be thought problematic, since the best way to formally spell out Hume's dictum itself appeals to the concept of intrinsicness.[43]

In this section I attempt to make trouble for combinatorial analyses of intrinsicality where Weatherson senses it may lurk, and to show that this poses a problem for anyone who, like Lewis, wants to make appeal to Com as capturing the Humean denial of necessary connections. Lewis' claim is that for any thing x, there is a world in which everything that exists is a part of a duplicate of x; that is

LD: $\forall x \exists w \exists y (Iyw \wedge Dyx \wedge \forall z (Izw \rightarrow z < y))$

Where 'Dyx' is to be read as 'y is a duplicate of x' and 'x<y' is to be read as 'x is a part of y' understood such that everything is a part of itself.

LD is entailed by Lewis' principle of recombination, seen above

I will argue that Com, and by implication LD, can only be justified by appeal to some combinatorial analysis of intrinsicality, which can in turn only be justified by appeal to Com, which creates a vicious circularity. A combinatorial analysis of intrinsicality is any which attempts to analyse the intrinsic properties as those which an object has independently of accompaniment, where the notion of independence here is modal. For example, consider the combinatorial analysis of intrinsicality proposed by Langton and Lewis.[44] Langton and Lewis firstly define what it is for a property to be independent of accompaniment. P is independent of accompaniment iff

1. It is possible for a lonely object to have P
2. It is possible for an accompanied object to have P
3. It is possible for a lonely object to lack P
4. It is possible for an accompanied object to lack P

[43] Weatherson, Brian, 'Intrinsic vs. Extrinsic Properties', Stanford Encyclopedia of Philosophy.
[44] [76]

Where an object is accompanied iff it is not lonely.

Langton and Lewis then define a basic intrinsic property as one which is (i) non-disjunctive (ii) not the negation of a disjunctive property and (iii) independent of accompaniment, where a property is disjunctive iff it can be expressed as a disjunction of (a conjunction of) natural properties but is not itself natural. The next notion to be defined is that of a duplicate: x and y are duplicates iff they share all and only the same basic intrinsic properties. The analysis of intrinsicality that Langton and Lewis then propose is that P is intrinsic iff duplicates never differ with respect to P.

Now the particular details of the Langton/Lewis account are not my concern here; my concern is with any such proposal that relies on the ability of some properties to be had independently of accompaniment. The idea behind this is that the intrinsic properties are such that whether or not a thing has them does not depend on what goes on in the world outside of that thing. As Yablo says, an intrinsic property is "a property a thing has (or lacks) regardless of what may be going on outside of itself."[45] A fortiori, then, whether or not a thing has or lacks an intrinsic property does not depend on what else exists in the world around it, and this leads to the thought behind the combinatorial analyses that the possibility of something having an intrinsic property should not depend on whether or not the thing is lonely or accompanied. This thought presupposes the truth of the Humean denial of necessary connections, since obviously properties can only be had independently of accompaniment if it is possible that *things* can exist independently of accompaniment to have those properties; and as we saw, the Humean principle relies on the notion of intrinsicality since it makes use of the concept of a duplicate.

That is where I think circularity lies; but let me be very clear at the outset what the objection is *not*: I am not objecting to the fact that we need to use the notion of intrinsicness in stating Com and that we need to use Com in our analysis of intrinsicality; *that* circularity is, I think, unobjectionable (I will say more about this below). The circularity I am objecting to is not the kind of circularity we get when two notions are defined or explicated in terms of one another. The circularity I am objecting to is epistemological; it is when two doctrines are such that warrant for either of them presupposes warrant for the other. In such a situation, it seems to me that

[45] [151, p479]

warrant for either doctrine is unobtainable; justification for either doctrine can never get off the ground. It is this that I will argue is the case concerning Com and intrinsicality. Justification for Com presupposes justification for a combinatorial analysis of intrinsicality, which in turn presupposes justification for Com; hence, neither can be justified.

What is obvious is that Com is only true if certain propositions concerning intrinsicality are true. Suppose, for example, that a thing a is intrinsically such that it bears a relation R to some thing b which is wholly distinct from a, or what is weaker, that it bears a relation to some thing or other which is wholly distinct from it. If there are such intrinsic properties then Com, and LD, are false; there will be no world in which there exists a lonely duplicate of a, since any duplicate of a must coexist with the thing it is R-related to. If Com is true, then there must be no such intrinsic properties. So the truth of Com presupposes the falsity of any analysis of intrinsicality which allows for such intrinsic properties. The conclusion I think we are led to is that we must have justification for an analysis of intrinsicality which rules out such properties as being intrinsic if we are to have justification for Com. Likewise, Com presupposes that every intrinsic property can be had by an unaccompanied object; this suggests that in order to be justified in believing Com we must have justification for accepting an analysis of intrinsicality which does not count as intrinsic properties which can only be had by things which are accompanied (e.g. properties which can only be had by universals, such as being composed of simpler universals, or properties which can only be had by tropes). This is why I claim that justification for Com presupposes justification for some combinatorial analysis of intrinsicality.

I claim that justification for a combinatorial analysis of intrinsicality is necessary for justification for Com. One might be worried that this is too strong. Perhaps justification for certain *applications* of the concept of intrinsicality is enough; and I can apply a concept I understand reliably without having justification for the analysis of the concept. Well it is true that I can be justified in making *some* applications of a concept without having justification for the analysis of the concept. I grant that one can justifiably hold that being square is intrinsic and that being the tallest tree in North Carolina which is not currently being observed by a man who is married to a woman older than 40 is not. A pre-theoretic grasp of the concept of intrinsicality is all that is necessary to justifiably make these claims. But not *all* applications of the concept of intrinsicality

can be justified solely on the basis of one's pre-theoretic grasp of the concept. There are some claims about intrinsicality which are not settled either way by pre-theoretic competence with the concept, claims which we expect to be settled by a successful analysis. And it is precisely these claims that the truth of Com depends on.

Our pre-theoretic usage of the concept of intrinsicality does not settle whether or not all intrinsic properties could be had by a lonely object, for the simple and obvious reason that our pre-theoretic intuitions concerning intrinsicality do not determine that there can be unaccompanied objects and yet they do determine that there are intrinsic properties. So a pre-theoretic grasp of the concept of intrinsicality cannot justify the claim that intrinsic properties can be had unaccompanied: we would only be warranted in making such a claim if we were warranted in holding some analysis of intrinsicality which entailed this claim. So the applications of the concept of intrinsicality that I need to be justified in making in order to be justified in holding Com are among those cases which I don't have a pre-theoretic justification for, those cases which I expect to be settled by an analysis of the concept. I stick to my claim, then: justification for the analysis of intrinsicality is necessary for justification for Com. One might object that Com is more basic than any claim concerning intrinsicality; that I simply do not need justification for Com and the fact that Com presupposes certain truths concerning intrinsicality simply limits what can count as an acceptable analysis of intrinsicality. That was perhaps Lewis' thought, since he didn't seem to think that Com was something that needed argument. But I think that response is misguided. No matter how entrenched or basic or fundamental you think the Humean denial of necessary connections is, that does not give you any right to claim that Com is so entrenched/basic/fundamental. The question that is at issue is precisely whether Com does indeed successfully capture the Humean doctrine. It only does so if certain propositions concerning intrinsicality are true, so we only have reason to accept Com if we have reason to accept these propositions. It is no good to claim that Com should be taken for granted since it is the Humean doctrine and that is not in need of further justification; that will convince no one since whether or not it captures the Humean doctrine is partly what we are currently trying to answer. To put what I think is ultimately the same point somewhat differently: duplication is a technical term, and we do not seem to have any reason to suppose that a duplicate of me is any more capable

of existing unaccompanied than I am until we have been given an analysis of intrinsicality (or of duplication) which implies this result; so we have no reason to suppose it is true that there can exist a lonely duplicate of any thing unless we have reason to accept the analysis of intrinsicality that tells us that is so.

I have argued that justification for the truth of Com requires justification for some combinatorial analysis of intrinsicality. What principle has explanatory priority: the principle which says what intrinsicality is or Com? Lewis is under strong pressure, I think, to say the latter. It is Com that, according to him, we rely on for most of our modal knowledge. But modal knowledge is required to check whether a combinatorial analysis of intrinsicality is correct, because the analysis is only acceptable if, by and large, the properties it tells us are intrinsic are the properties we pre-theoretically thought were intrinsic. There is some room for divergence of course; an analysis can inform us by settling cases and so on. But if there is too much divergence then we have no justification for thinking that the concept we have analysed is the concept we use the English term 'intrinsic' to signify. So in order to know that an analysis of intrinsicality is right, we need to know that it gives pre-theoretically sound results in most cases; which involves us knowing what properties could be had and lacked by lonely and accompanied objects. So knowledge that the analysis of intrinsicality is correct requires modal knowledge, which, according to Lewis, we by and large obtain by (tacit) application of Com. But doesn't that suggest that our warrant for Com should be independent of, or prior to, the truth of the analysis of intrinsicality? Now of course, Lewis does not say that Com is our *only* route to modal knowledge, only that it gives us our "best handle" on what is possible. It would be perfectly consistent for him to claim that it is some other principle, one that does not make use of the concept of intrinsicality, that we use in order to justify the analysis of intrinsicality, which we in turn use to justify Com which in turn lets us get a lot more modal knowledge. But the modal knowledge that is needed to justify the analysis of intrinsicality includes the knowledge that things can exist unaccompanied: and isn't that precisely the knowledge that we are meant to obtain through application of Com?

The problem can be stated as follows: either Lewis relied (tacitly or explicitly) on Com to make the modal claims necessary to justify his analysis of intrinsicality or he didn't. Assume he did: in that case we have no more reason to accept the analysis of intrinsicality than we do to

accept Com; in which case we should doubt both since the truth of Com is what is up for debate. If I don't currently believe either of two claims, and if acquiring justification for one requires justification for the other, then I will not be moved to accept either. Assume then that Lewis didn't rely on Com, either tacitly or explicitly: then how did he get the knowledge necessary to justify the analysis that things could exist unaccompanied? Isn't the only reason we have to suppose that things could exist unaccompanied that Com is true? So the worry is that if Lewis did not rely on Com then he will be a counter-example to his own claim that Com underpins this sort of modal knowledge.

What exactly does this show? Circularity in analysis is not always bad. Consider Quine's critique of analyticity.[46] Quine argued that the analysis of analyticity was circular because it relied on the concept of synonymy which, in turn, relied on the concept of necessity, which relied on the concept of analyticity. Now of course one can attempt to break this circle; most contemporary philosophers are disinclined to analyse necessity in terms of analyticity. But even if you accept each analysis, it does not seem like Quine has given good reason to give up on analyticity. For Quine has not shown yet that there is an *epistemic* circularity; that one cannot come to a proper understanding of what analyticity is. That would only follow if in order to grasp each of the concepts involved one would have to have prior grasp of the analysis. But we seemingly acquire a grasp of synonymy without grasping its definition. Competence as a language user requires the ability to tell, more or less, when words we know the meanings of mean the same as other words we know the meaning of. In that case grasp of synonymy allows us to grasp analyticity; it is no matter that the analysis of synonymy makes appeal to the concept of analyticity, because our understanding of synonymy does not rest on understanding of this analysis. The lesson we should take from the Quinean example is that circularity in the analyses of a cluster of concepts does not mean that we cannot acquire a grasp of the concepts in that cluster. If there is one of those concepts that can be grasped without understanding its analysis then that gives us means to thereby grasp all the concepts in that cluster.

But this does not help in the above case concerning recombination and intrinsicality. Grant for the sake of argument that one can grasp what it is to be intrinsic without understanding

[46][115]

the analysis of intrinsicality. That means that I can come to have an *understanding* of the principle of recombination. But this is not the goal. The goal is to gain *justification* for Com. Understanding the principle is not sufficient to obtain that justification (it is not self-evident, after all). The circularity is vicious not because it results in our not being able to grasp the concept of intrinsicality or the principle of recombination, but because the truth of Com depends on the success of the analysis of intrinsicality. This in turn depends on the space of possibilities being a certain way; which, according to Lewis, I check by tacit application of Com. But if it is tacit grasp of Com that I use to infer what is possible, then my claims about what is possible are only reliable if Com is in fact true. But that was what was up for question in the first instance.

So the worry is not about how we acquire a grasp of Com or intrinsicality. Circularity worries would only be a problem there if grasp of the analyses in question were necessary for grasp of the concepts. Rather, the worry grants our understanding of recombination and intrinsicality, but shows that the truth of Com presupposes certain claims about intrinsicality — namely, that intrinsic properties can be had by unaccompanied objects. And either those claims about intrinsicality are justified by presupposing Com, in which case the justification is circular and we have no reason to believe either Com or the claims about intrinsicality, or the claims about intrinsicality are justified without presupposing Com, in which case the need for Com is undermined, since the point of it was in part to secure the very claim in question — that there can be unaccompanied objects.

What conclusion should we draw? Lewis claims that the principle of recombination is the principle that informs us, by and large, as to what is possible; and his reason for thinking this is that it captures the Humean denial of necessary connections, which he takes to give us our best handle on what is possible. To make good these two claims he needs to convince us of two things: (i) that there is a version of the principle of recombination that entails LD, since LD is what he takes to be the denial of necessary connections, and (ii) that such a version of the principle is true, since we should not believe that we rely tacitly on a *false* principle when making our modal inferences.[47] The problem arises because in order to convince us of (ii) Lewis needs to make certain modal claims, since that is necessary to justify the claims concerning intrinsicality that

[47]That's not to say that we *don't* rely on a false principle, of course; nor to say that any principle we rely on cannot go wrong. The claim I am making concerns warrant and not truth.

(ii) relies on. And if those modal claims relied on the version of recombination in question then they presuppose the truth of (ii) which they were supposed to establish. And if they did not rely on that principle then this goes against the hypothesis that recombination is the principle that we use to infer the possibility of the situations in question.

Lewis is in a bind. He needs a modal analysis of intrinsicality since only a modal analysis will have a chance of justifying the claim that duplicates of objects can exist unaccompanied; but he needs the modal analysis to be justifiable without presupposing those very claims concerning duplicates. The latter seems impossible, however; for the modal analysis relies on the truth of the Humean principle which assumes that duplicates can exist unaccompanied. I see no way out of this circle, and hence no way to justify Com or LD.

End of Digression

But even if we could obtain a warranted belief that Com is true, this doesn't seem to be a problem for the truthmaker theorist per se; for at least some theories of truthmakers are compatible with Com. Hence, by Lewis' own lights, they are compatible with the Humean denial of necessary connections. So where, according to Lewis, is the problem? For example, consider the view that truthmakers are tropes. So the truthmaker for [A is F] is the trope of A's being F. Lewis objects that there is a necessary connection because the existence of the trope entails the existence of A. But there need be no violation of Com. The trope theorist can agree that there can exist a *duplicate* of A's being F without there existing either (a counterpart of) A or a duplicate of A. Indeed, depending on the details of their theory, they can agree that it is possible that a duplicate of the trope exist without any other thing existing at all. For while it is an essential property of the trope (we are supposing) that it belongs to A, it is not an intrinsic property of the trope. So there can exist a lonely duplicate of the trope, which is not a counterpart of/identical to the trope itself; and so we have the existence of the trope necessitating the truth of [A is F] with no violation of Com; hence, by Lewis' own lights, no violation of the denial of necessary connections. So Com might be incompatible with *some* theories of truthmaking. But it is not incompatible with them all; and so if Lewis is right that Com captures the Humean denial of necessary connections then he is wrong that the denial of necessary connections is incompatible

with truthmaker theory.

Furthermore, we still have a problem with set theory: if it's impossible for there to be unaccompanied things, because there must also be the resulting impure sets, then a fortiori it is impossible for there to be unaccompanied duplicates of things. So it seems hopeless to try and capture the thought behind DNC using talk of duplicates. Perhaps, however, we could capture the thought behind HM by making appeal to duplicates. There are three options to formulate HM in such a way.

HMD^1: if a≠b then there could exist a duplicate of a in a world such that any duplicate of b is a part of the duplicate of a, or vice-versa.[48]

HMD^2: if a≠b then there could exist a duplicate of a in a world where (a counterpart of) b does not exist.

HMD^3: if a≠b then (a counterpart of) a could exist in a world in which there is no duplicate of b.

Translated into counterpart theory, where 'Dxy' is read as 'x is a duplicate of y', these read:

$$HMD^1 : a \neq b \rightarrow \exists w \exists x (Ixw \wedge Dxa \wedge \forall y((Dyb \wedge Iyw) \rightarrow (x < y \vee y < x)))$$

$$HMD^2 : a \neq b \rightarrow \exists w \exists x (Dxa \wedge Ixw \wedge \neg \exists y (Cyb \wedge Iyw))$$

$$HMD^3 : a \neq b \rightarrow \exists w \exists x (Cxa \wedge Ixw \wedge \neg \exists y (Dyb \wedge Iyw))$$

But we should not accept any of HMD^1, HMD^2 and HMD^3, as counter-examples are available to each, at least if the following plausible principle is true: that the intrinsic properties of a set supervene on the intrinsic properties of its members. That sounds intuitively plausible: that all one needs to do to fix the intrinsic properties of a set are to fix the intrinsic properties of what is in that set. As a first attempt at making this idea precise let us try: If there is a difference in the intrinsic properties of a set S and a set S* then one of those sets contains something which differs in its intrinsic properties from some member of the other set. In other words, whenever we have

[48] We can't claim that a duplicate of a could exist without a duplicate of b due to the Castor and Pollux example. Since they are intrinsically identical, there cannot exist a duplicate of one without there existing a duplicate of the other; but intuitively this is no problem for the denier of necessary connections. C.f. the discussion of HM^*_{CT} above.

two sets which are not duplicates we can find a pair of objects x, y such that x belongs to one set and y to the other and that x and y are not duplicates. But this can't quite be right because that would make the singleton of A and the set containing two duplicates of A duplicates. But surely it is intrinsic to a set how many members it has. To fix this we need to build in the further condition that duplicate sets be equinumerous. So let our principle be that if two sets S and S* differ in their intrinsic properties then either (i) S and S* are not equinumerous or (ii) we can find a pair of objects x, y such that x belongs to S and y to S* and such that x and y are not duplicates. This principle will do, I think.

There are some who will reject the principle. Tom Baldwin, for example, thinks that the singleton of a thing A and the singleton of a duplicate of A which is not A, call it A*, provide a counter-example to this claim, because he thinks that it is intrinsic to a set what members it has but not intrinsic to things that they are members of the sets they are members of.[49] If that is true then there is a difference in the intrinsic properties of {A} and {A*} because {A} but not {A*} has the intrinsic property of having A as a member. But I see no reason to treat 'having A as a member' as intrinsic and 'being a member of singleton A' as extrinsic; the properties should either both be intrinsic or extrinsic, and once this is admitted we no longer have a counter-example. Either A* is not a duplicate of A because A but not A* has the intrinsic property of being a member of {A} or, what is more likely, {A} and {A*} are duplicates because the only properties they differ in are the external properties of having A or having A* as a member. With this counter-example dispatched, I see no way to cause trouble for the principle.

If the above principle is true then HMD^1 is false. The existence of a thing is sufficient for the existence of its singleton; a fortiori the existence of a duplicate of a thing is sufficient for the existence of the singleton of that duplicate. In which case, since singletons are equinumerous with one another, in any world in which there is a duplicate of A there is also a duplicate of the wholly distinct thing singleton A, making A and its singleton counter-examples to HMD^1.

HMD^2 will fail if there are any objects such that it is sufficient for them to be that very object that they have the intrinsic properties they in fact have. For suppose there is some such object B. Necessarily if there is a duplicate of B, B exists, since that duplicate would, ex hypothesi, be B.

[49] See [13].

But B's existence is sufficient for the existence of singleton B, so any world in which a duplicate of B exists is a world in which singleton B exists, making B and singleton B counter-examples to HMD^2. Are there things such that it is sufficient for something to be that very thing that it be a duplicate of that thing? Well, the empty set and universals appear to be good candidates. And again, I don't particularly want to claim that these things exist, but I don't want to be told they don't exist because that would violate HMD^2.

Similarly, HMD^3 will be false if there are any objects such that it is a necessary condition for their existence that they have the intrinsic properties they in fact have. For consider some such object, call it C. Necessarily if C exists, a duplicate of C exists, since C would be that duplicate. But the existence of C is sufficient for the existence of singleton C which, given the principle (defended above) that the intrinsic properties of a set supervene on the intrinsic properties of its members, is also such that it has its intrinsic properties necessarily. This makes C and singleton C counter-examples to HMD^3, for C cannot exist without a duplicate of singleton C existing. Are there objects such that their intrinsic properties are essential to them? Again, the empty set and universals seem to be likely contenders, and also tropes. In which case the lack of incoherence surrounding the hypothesis that such things exist is reason not to accept HMD^3.

4.2.5 Necessary Connections and Ontological Dependence

Things are not looking good for the denial of necessary connections; but perhaps all is not lost. Notice that there seems to be an explanation for the necessary connection between a singleton and its member that would be missing in the case of a supposed necessary connection between you and me; namely that in the former case but not the latter there is a relation of ontological priority between the things in question. Ontological priority, as I will use the term, is the converse of Lowe's relation of identity dependence.[50]

x is identity dependent on y iff what it is that makes x the very thing x is in part that the thing y is the very thing y. Sets, for example, are identity dependent on their members. What makes a set S the very set S and not some other thing is that it has as its members the very members that it has. The identity conditions for sets are just that they have the exact members

[50][83, p147]

they have. A change in members entails a change in set. Similarly, tropes are identity dependent on the things of which they are properties. The redness of this apple is identity dependent on the apple, because what it is for the trope to be that very trope is that the apple is that very apple. If the apple was a different apple, then the trope would be a different trope. If x is identity dependent on y, then y is ontologically prior to x. A set is identity dependent on its members; since what it is to be the very set S is that the members of S are the very things they are; hence the members are ontologically prior to the set.

Now if a is ontologically prior to b — if the identity of b depends on the identity of a — then we are committed to a necessary connection between b and a. The fact that b is the very thing it is depends on the fact that a is the very thing it is. If b exists, then it is the very thing b; it wouldn't be the very thing b if a were not the very thing a. a is only the very thing a if it exists. So for b to exist, a must exist.

If this line of reasoning is correct, we get the following principle.

OP: If A is ontologically prior to B then, necessarily, if B exists, then A exists.

In symbols.

$$OP : Rab \rightarrow \Box(\exists x(x = b) \rightarrow \exists y(y = a)))$$

Where 'Rab' is read as 'a is ontologically prior to b'.

What would be nice is if the principle ran both ways; if as well as it being a sufficient condition for there to be a necessary connection between B's existence and A's existence that B was identity dependent on A, that it was also a necessary condition. That is

$$OP^* : Rab \leftrightarrow \Box(\exists x(x = b) \rightarrow \exists y(y = a)))$$

For then we could infer that it is possible for a to exist without b on the grounds that a is not identity dependent on b.

But OP*, on the face of it, has counterexamples. The existence of a set depends on the existence of its members; but also, as I said above, it seems that the existence of the members is sufficient for the existence of the set. Necessarily if the members of S exist, then S exists; but

the members of S are not identity dependent on S — the members are ontologically prior to the set, not vice-versa — so this necessary connection would be a counter-example to OP*.

What is interesting, though, is that one can accept OP* and resist this counter-example if one adopts counterpart theory. Consider Fine's contention that one cannot give a modal analysis of essence, because essence is asymmetric in a way necessary dependence is not.[51] In particular, he says that it is in the essence of singleton Socrates that it has Socrates as a member, but not in the essence of Socrates that he is a member of singleton Socrates, even though necessarily if Socrates exists he is a member of singleton Socrates. The counterpart theorist can account for this asymmetry. The counterpart theorist can define 'A is essentially F' as 'Necessarily if A exists, A is F' and also, if they wish, hold on to Fine's claim that Socrates is not essentially a member of singleton Socrates. Why? Because the set of the counterparts of the members of S need not be a counterpart of S.

The counterpart theorist can agree that singleton Socrates essentially has Socrates as a member. They will analyse this as: every counterpart of singleton Socrates has as a member a counterpart of Socrates. But they are not, in saying this, forced into saying that every singleton of a counterpart of Socrates is a counterpart of singleton Socrates; hence they can reject, with Fine, the claim that Socrates is essentially a member of singleton Socrates. Here is one very easy way they can do this (but not the only way). They can claim that the only counterpart of a set S is S. That does not seem unwarranted; one might think it is the best way to make sense of the intuitive thought that a set is essentially exactly the way it in fact is. But no one would want to claim that the only counterpart of Socrates is Socrates. Socrates has some accidental properties; in particular he has a counterpart who is not snub-nosed; call him SOC. SOC is a counterpart of Socrates but singleton SOC is not a counterpart of singleton Socrates. The only counterpart of singleton Socrates is itself. So while it is necessary that if singleton Socrates exists it has Socrates as a member, Socrates can exist without being a member of singleton Socrates. (This is of course to deny Lewis' view in Parts of Classes where he holds that the singleton {a} is a counterpart of the singleton {b} iff a is a counterpart of b.)[52]

This claim — that the counterpart of a set S needn't be the set of the counterparts of the

[51] [38]
[52] [67, p37].

members of S — has some odd consequences. Crispin Wright has objected that it yields the result that objects could have had different singletons, which he finds absurd. Fair enough; I admit it is a cost, but it is not obviously a reductio of the position. My argument is not that I think the counterpart theorist *should* hold OP*; I am claiming only that they *can*. Personally, I think they shouldn't: it seems very plausible to hold (which is an extension of Lewis' view[53]) that a set S* is a counterpart of a set S iff (i) S and S* are equinumerous[54] and (ii) there is a bijection on S and S* such that every member of S is linked to a member of S* which is its counterpart.

Still, the counterpart theorist can accept OP* (or rather, its counterpart theoretic translation) if she wishes, and so there is at least one principle that perhaps deserves to be called the denial of necessary connections that Lewis can, if he wants, make claim to that his opponents cannot. But again, this at most puts *constraints* on what truthmakers could be; it does not rule out truthmaker theory altogether. In particular, OP* is perfectly compatible with the view that truthmakers are tropes inhering in substrata, since tropes are identity dependent on the substance to which they belong. What makes the trope of A's F-ness the very trope it is is, in part, that it belongs to the very thing A; so the identity of the trope depends on the identity of the substance, and so the substance is ontologically prior to the trope. Likewise with the view that states of affairs are truthmakers; states of affairs are identity dependent on the thin particulars which constitute them and so, if OP* is the principle we must adhere to, there is no problem in the necessary connection that obtains between the state of affairs and the thin particular. It seems then that OP*, even if it is true, poses no real problem for the truthmaker theorist.

4.3 Modal Truthmakers

Having defended a version of truthmaker theory, now we can turn to the question that presently interests us: What are the truthmakers for modal truths? By this I mean, what are the truthmakers for truths of the form 'p is necessary/contingent' as opposed to what are the truthmakers for necessary truths. (For I would reject the demand for a truthmaker for [there is no married

[53][ibid.]

[54]This is to avoid the absurdity that singletons could have had more than one member

bachelor], but nothing we have said so far lets us rule out the demand for a truthmaker for [necessarily, there is no married bachelor].) However, we can take as a lesson from the discussion on truthmaker maximalism above that we do not need to find truthmakers for possibility claims *and* necessity claims. Since necessity is simply the absence of the possibility of the negation, and possibility the absence of necessity of the negation, it will suffice to locate truthmakers for possibility claims or for necessity claims. If we find truthmakers for possibility claims then the source of necessity will simply be the lack of a truthmaker for the possibility of the negation, and vice-versa for the source of possibility if we find truthmakers for necessity claims. The only person I know of who has given real effort to finding truthmakers for modal truths given an actualist ontology is Armstrong, so we shall look at his views now.

Firstly, let us consider Armstrong on possibility claims. Armstrong offers a proof that for every true claim of mere possibility there is a truthmaker for that truth to be found in actuality.[55]

1. a makes it true that p (Assumption)
2. p is contingent (Assumption)
3. From 2, p entails that $\Diamond\neg p$
4. If a makes q true and q entails r then a makes r true (Assumption)
5. From 1, 3 and 4, a makes it true that $\Diamond\neg p$

If this argument is sound it establishes that what makes it true that p, for some contingent p, is also what makes it true that p could have been false, i.e. that it is possible that $\neg p$. Now clearly I have a problem with the assumption at 1, since I reject truthmaker maximalism; for me the most one could get from an argument like this is that *some* truths of mere possibility have actual truthmakers. The truth that there could be unicorns, however, will be left unaccounted for since I reject the claim that there is a truthmaker for the fact that there are no unicorns. Likewise, Armstrong uses this argument to account for the possibility of alien entities (particulars, properties and relations that are not combinatorially constructible from actual entities) by noticing that the truthmaker for the fact that all the actual things are *all* the actual things will

[55] See [9, p84]; I have changed the presentation of the argument, but it remains the same as Armstrong's in essentials.

be a truthmaker for the possibility of aliens.[56] But this move, of course, requires the existence of totality facts, which I have rejected above.

But even granting truthmaker maximalism this argument does not seem to work. It is step 3 that is worrisome. Armstrong says that 3 follows from 2 and "the nature of the contingency of propositions."[57] Well it certainly follows from the fact that p is contingent that ¬p is possible, but that is not what is being said at 3; what is being said is that it follows from p — the proposition itself, rather than the fact that it is contingent — that ¬p is possible. And it is necessary that it is this claim that is being made, of course, for the application of the entailment principle (4); it must be p that entails the possibility of ¬p, not simply the fact that p is contingent, since it is p, and not the fact that p is contingent, that we know has a truthmaker.

But does p entail that ¬p is possible? Armstrong says that "Given the attractive S5 modal system, if p is contingent, it is a necessary truth that it is contingent. This may help to quell any doubts one may have about step 3 in the argument."[58] But as I have already said, I am not prepared to accept S5 on faith; I want an account of the source of modal truth which allows for the possibility of propositions having their modal status as a matter of contingency. As such, I do not accept that p entails $\Diamond\neg$p, and so I do not accept Armstrong's proof of the existence of truthmakers for facts of mere possibility, even putting aside concerns about the assumption of truthmaker maximalism.

In fact, it is very odd that Armstrong thinks that an appeal to S5 can justify the move at 3. If S5 is true and entailment is classical then step 3 is justified; but then the above proof is not needed. From step 4 and the paradigmatic truthmaking relation we can prove that every thing is a truthmaker for truths of possibility, since truths of possibility are themselves necessary (in S5) and necessary truths are (classically) entailed by any set of formulae. Since every thing is a truthmaker for some proposition (namely the proposition that that thing exists) it follows that every thing makes truths of possibility true, since the proposition that the thing exists entails the proposition that such-and-such is possible. Armstrong, however, does not want the notion of entailment in the assumption at (4) to be classical, for he does not want to hold that necessary

[56] ibid. p86-89
[57] ibid. p84
[58] ibid. p84-85

truths are made true by every thing.[59] He thinks, and I agree, that this is too easy a solution to the problem of necessary truths to be adequate. Some kind of relevant entailment is required, then, in the formulation of the entailment principle at (4). But in that case, even if one believes S5 to be the correct system of modal logic, this will not "quell the doubts" about step 3, for p does not relevantly entail $\Diamond\neg p$, no matter whether or not $\Diamond\neg p$ is necessary. Armstrong cannot have his cake and eat it: if he wishes to appeal to relevant entailment then he is not justified in making the move from 2 to 3; he *is* justified in making this move if he wishes to appeal to classical entailment and S5, but at the cost of trivialising truthmakers for necessary truths.

Now in fact Armstrong had put forward a different version of the argument in an earlier paper which avoids this particular worry. That version ran as follows[60]:

1. a makes it true that p (Assumption)
2. a makes it true that p is contingent (Assumption)
3. From 1 and 2, a makes it true that (p and p is contingent)
4. (p and p is contingent) entails $\Diamond\neg p$
5. If a makes q true and q entails r then a makes r true (Assumption)
6. From 3, 4 and 5, a makes it true that $\Diamond\neg p$

Here 2 is not being said to follow from 1 but is an explicit assumption of the argument. Well, in this case I have no problem with the validity of the argument. Sure, if a makes it true that p is contingent, and if the entailment principle is true, then a makes it true that $\neg p$ is possible. But I question Armstrong's right to make the assumption at 2, for it seems simply to beg the question: if one doubts that the possibility of $\neg p$ obtains in virtue of a then it is not clear why one would be willing to accept that the *contingency* of p obtains in virtue of a. The contingency of a proposition *just is* the conjunction of its truth and the possibility of its negation. Now we know, *ex hypothesi*, that the truth of p obtains in virtue of a; in that case isn't the further claim

[59] ibid. p10-11

[60] [8, p15]; again, the presentation is changed from Armstrong, but not in any way that affects the argument. Armstrong, in the later work, thanks Marian David for offering a simplification of this argument, so I suppose it is he that is to blame for the introduction of the particular problem noted above.

that a also accounts for the contingency of p simply tantamount to claiming that a also suffices for the possibility of the negation of p? But that is precisely what the argument is aiming to prove, so it can't be presupposed by any premise without begging the question.

Armstrong offers the following rather obscure defence of premise 2[61]

> a is something in the world, some state of affairs or other entity depending on just what truthmakers are postulated, a matter that depends on one's whole metaphysics. Whatever a is, in the cases we are considering it is a contingent being. Could the contingency of a lie outside a? It does not seem possible. It cannot be a relation that a has to something beyond itself. So a is the truthmaker for the proposition [p is contingent].

This last "so" is completely beyond me! Certainly Armstrong must hold that a is a contingent being, given the contingency of p and truthmaker necessitarianism. The next point seems to be that a is the truthmaker for its own contingency: that is, that [a might not have existed] is true in virtue of a. Well that's not obvious, but even granting that a contingent being makes true the proposition that it might not have existed, how are we supposed to go from this to the claim that a is the truthmaker for $\Diamond\neg p$? It does not follow from the fact that the truthmaker for p might not have existed that p is contingent, for it may be necessary that there exists *some* truthmaker for p even if it is not necessary that the actual truthmaker for p exists. (Armstrong is not committed to, nor should one accept, the claim that if p is made true by a then necessarily if p is true it is made true by a. [There is a human] is made true by many individuals, all of whom might not have existed and that proposition still have been made true.) So even if a makes it true both that p and that a might not have existed, we do not seem to have anything to suggest that a makes it true that ¬p is possible; for all that has been said so far it may be that in all the worlds in which a does not exist there is some other thing that makes it true that p. And so the argument for premise 2 is unconvincing; all Armstrong can do is assume premise 2, which begs the question.

I think we should reject the above 'proofs' that there are truthmakers for truths of mere pos-

[61] ibid. I have changed occurrences of 'T', which is what Armstrong names the truthmaker for p, to 'a'; I have not used 'T' because I want to avoid any possible confusion with a truth predicate.

sibility amongst the actual. Armstrong, however, goes on to offer a story about what the truths for mere possibility are and we must now look at that. Consider the possibly true proposition [there is a unicorn]. Armstrong aims to account for this possibility combinatorially; what makes [possibly there are unicorns] true, says Armstrong, just are actually existing things that are combinable to make a unicorn: say an actual horse and an actual horn. Of course we also need the truth that these actual things are combinable, since some actual entities are not, such as roundness and squareness. So what makes it true that these particular things — this horse and this horn — are combinable? No more than the horse and the unicorn themselves, says Armstrong, since their combinability is necessitated by their nature (while the non-combinability of roundness and squareness is necessitated by their nature).[62]

I have two problems with this. Firstly, the possibility of aliens is left unaccounted for; by definition aliens are not the result of any combination of actual entities and so, since we have rejected Armstrong's suggestion that the truthmakers for claims concerning the possibility of aliens are totality facts concerning all the things there in fact are, this possibility is left ungrounded. Secondly, and more importantly, we now have an appeal to natures, and for me this is putting the cart before the horse. The horse and the horn are combinable *because* unicorns are possible, not the other way round. The modal properties of a thing, as I argued in chapter 3, are determined by what properties are possibly instantiated, so the fact that the horse has the property of being combinable with the horn (or vice-versa) is true *because* there is, in logical space, a unicorn. Armstrong has got the direction of explanation the wrong way round: facts about essence are explained by the facts concerning how things could have been, whereas he is trying to explain how things could have been by appealing to facts about essence.

Armstrong's story about truthmakers for claims of mere possibility is unsuccessful, I claim. But this is not a problem if his story about truthmakers for claims of necessity works, for then we could simply characterise the possibilities as those propositions whose negation is not made necessary. Unfortunately, instead of focussing on truths of the form [Necessarily, p], Armstrong turns his attention to truths of the form p which happen to be necessary. When discussing possibility he considers propositions such as [possibly, there are unicorns] but when discussing

[62] [9, p91-93]

necessity he considers propositions such as [7+5=12] when he should be considering propositions such as [Necessarily, 7+5=12]. But let us look at what he thinks the truthmaker for [7+5=12] is, since it may shed light on the truthmaker for [Necessarily, 7+5=12].

Armstrong thinks that the truthmaker for [7+5=12] is simply the numbers involved — 7, 5 and 12 — if they exist. Why? Because "Given the entities 7, 5 and 12, then they must, necessarily must, be related in this way [I.e. such that when you add the first two you get the last]. . . So (by the entailment principle) truthmakers for the existence of the entities should be a sufficient truthmaker for the necessary truth."[63] Armstrong thinks that whatever makes it true that the numbers exist also makes true the truths of mathematics.

Let us grant that 7, 5 and 12 make true [7+5=12]; it seems to me quite plausible. Do they make true [Necessarily, 7+5=12]? I think not. We face the question as to whether numbers exist necessarily or contingently. If they exist necessarily then we have an account of the truth of [Necessarily, 7+5=12], but at the cost of introducing new necessary truths that have not been grounded. For what now is the truthmaker for [Necessarily the number 7 exists]? Not the number 7 presumably: that only makes true the proposition that 7 exists, not that it *necessarily* exists.

Armstrong, however, wants to resist the claim that there are necessary existents. He claims that numbers exist contingently, but that this is compatible with the necessity of [7+5=12] because all that is required for the truth of that proposition is the *possibility* of the existence of the numbers involved; they need not in fact exist. In that case we need not appeal to any necessary truths concerning what exists to account for the truth of [Necessarily, 7+5=12], it will suffice for the truth of that proposition that necessarily it is *possible* that 7, 5 and 12 exist. But still, there is an ungrounded necessary truth: what makes true the proposition [Necessarily, there could be the numbers 7, 5 and 12]? Armstrong offers us no explanation.

So even if Armstrong has given an adequate story as to the truthmaker of [7+5=12] he has given us no explanation of the truth of [Necessarily, 7+5=12], which is where his attention should have been. Either the truth of [7+5=12] requires the actual existence of the numbers involved or it only requires their possible existence; but either way we are left with ungrounded necessary truths when we try to account for the necessity of [7+5=12]. If the truth of [7+5=12] requires

[63] ibid. p98-99

the actual existence of 7, 5 and 12 then the truth of [Necessarily, 7+5=12] requires that 7, 5 and 12 be necessary existents, and we have been given no truthmaker for the claims that these numbers exist necessarily. On the other hand, if the truth of [7+5=12] requires only the possible existence of 7, 5 and 12 then in order for [Necessarily, 7+5=12] to be true it must necessarily be possible that 7, 5 and 12 exist; but we have been given no truthmaker for the claim that such a situation is necessarily possible. (We have not even been given a truthmaker for the claim that such a situation is actually possible.)

I conclude that Armstrong's attempt to locate truthmakers for modal truths among actuality is unsuccessful. And on reflection this is not surprising, for it is deeply mysterious how actual entities could suffice for the truth of modal truths. Consider some contingently existing thing, a. What actual thing could make it true that a might not have existed? Not a, seemingly: a makes it true that a exists, but not that it could have failed to exist. Likewise with necessary existents, assuming there are *some*: what makes it true that 5, for example, necessarily exists? Again, not 5, since 5 suffices only for the truth of [5 exists] and not [Necessarily, 5 exists]. But how could something other than 5 make it true that 5 exists necessarily? Unless we make appeal to natures or essences here, and claim that it is in the essence of 5 that it exists necessarily, or the essence of a that it exists contingently, then actuality seems ill-equipped to account for these truths. But, as I have argued above, appeal to essence facts here is getting things the wrong way round. Facts about essence are settled by the modal facts, not vice-versa.

What actual thing makes it true that there could have been more things than there in fact are? How could any actual thing make that true? What actual thing makes it true that as a matter of necessity no proposition is both true and false? There is nothing in actuality, so far as I can see, that makes this true. I conclude that truthmakers for modal truths are unlikely to be found among the ontology of the actual. The actual things only make true what is in fact the case, not what must have been the case or what could have been the case but isn't. If we want truthmakers for modal truths, then, perhaps we ought to admit non-actual ontology.

Chapter 5

Modal Realism

We have looked at the theory of truthmakers and attempted to locate truthmakers for modal truths among the ontology of the actual, and I concluded that this was unlikely to succeed. Perhaps, then, if modal truth is to be grounded in ontology we must admit of the merely possible in addition to the actual. This would be a response to the problem of modal truth analogous to the eternalists' response to the problem of temporal truth. Just as the eternalist argues for the existence of non-present things to ground the truth of facts concerning how the world was or will be, perhaps we should admit the existence of merely possible things to ground the truth of facts concerning how the world could be.

This is Lewis' response to the problem of modal truth, and in this chapter I will examine Lewis' Modal Realism. Lewis claims that the actual world is one among many; that for any way a world *could* be, there is a world that is that way — a world causally unconnected from any other world, and not related to any other by any spatial or temporal measure.[1] So while we previously supposed that there were no talking donkeys or flying pigs, Lewis tells us that in fact there are such things. Such things are not *actual*; but that only means that they do not exist at *our* world, 'actual' being an indexical term like 'here' or 'now' which picks out the world of the utterer. According to Lewis, our intuitions that there are no talking donkeys arise because ordinarily our quantifier is restricted to what is actual: when we say 'there are

[1] [66]

no talking donkeys', ordinarily this has the truth-conditions that there are no actual talking donkeys, which is true; but unrestrictedly it is, according to Lewis, false to claim that there are no talking donkeys, for at some world there are such things. This theory is meant to afford us many benefits; of particular interest to us are the benefits claimed concerning modality.

Since Lewisian realism reduces talk of the modal to talk of non-actual individuals, it locates the source of the modal truth in non-actual ontology. It is possible that there is a talking donkey in virtue of the fact that there is (unrestrictedly speaking) a talking donkey; it is necessary that there are no round squares because there are no (unrestrictedly speaking) round squares. There are non-actual truthmakers for truths of possibility then. The possibility of p holds in virtue of the existence of a world at which p; the necessity of q holds, on the other hand, in virtue of the fact that there is no truthmaker for $\Diamond\neg q$ — i.e. that there is no world at which q is false. So the necessary truths are negative facts on this account; they themselves do not have truthmakers, but nevertheless their truth is not left ungrounded: they are true because the corresponding claims of possibility lack truthmakers. Do the truthmakers for the claims of possibility — namely, the worlds[2] — necessitate the truth of those possibility claims? Yes, because the equivalence between what is true at a world and what is possible is itself a necessary truth if it is true at all; and so in every world in which it is true that there is a world at which p it is possible that p.[3]

Armstrong objects to the Lewisian account of truthmakers for modal truths on the grounds that it gets the direction of explanation wrong.[4] According to Lewisian realism it is necessary that p because p is true at every world whereas Armstrong thinks that even if one believes in the Lewis pluriverse one should hold that p is true at every world because it is necessary. I confess I do not see the worry. Isn't the Lewisian reading of the biconditional simply a consequence of the fact that they are offering a *reduction* of the modal? To object to the reduction on the grounds that it forces a reductive direction of explanation on the relevant biconditional is just to object to it on the grounds that it is a reduction. That is hardly going to convince one's opponent.

[2]Worlds will always suffice as truthmakers for possibility claims, but they will not usually be minimal truthmakers; for example, while a world w containing a talking donkey is a truthmaker for [there could be a talking donkey], the talking donkey itself is the minimal truthmaker. Worlds will only be minimal truthmakers for claims concerning the possibility of some maximally specified situation.

[3]There is an issue as to how to evaluate claims about the plurality of worlds at a world. Such claims are called advanced modalizing by John Divers, and I will treat them in the manner he suggests: namely that truths concerning the plurality of worlds are themselves taken to be necessary at every world. See [31].

[4][9, p95-96]

If the Lewisian is correct about what there is then there are non-actual truthmakers for modal truths, but should we accept Lewisian realism? I shall look at four objections to the theory; the first two I shall conclude fail; but the last two, which are related to each other and also in a way to the second objection, I argue succeed.

5.1 Metaphysical Nihilism

The first objection to Lewisian realism I will consider is that it does not provide an adequate account of the modal because it does not account for every possibility. In particular, so the objection goes, it does not account for the possibility of there being an empty world — a world in which there are no concrete objects. The incompatibility of such a world and Lewisian realism has been argued for by Gonzalo Rodriguez-Pereyra.[5] His reasoning is as follows. Assume there could be no concrete objects. That is to say that there is a world where nothing concrete exists; i.e. there is a maximal mereological sum of spatio-temporally related individuals where nothing concrete exists. Such a sum is spatio-temporally located, and spatio-temporally located objects are, says Rodriguez-Pereyra, concrete objects. In that case there is a concrete object where no concrete objects exist. But that is a contradiction; "the world where nothing concrete exists is a world where some concrete objects exist!"[6]

What should we make of Rodriguez-Pereyra's argument? Certainly there is a step that the Lewisian realist could, perhaps should, resist. That is the step from 'such and such is a sum of spatio-temporally related individuals' to 'such and such is a concrete individual'. Must we agree with Rodriguez-Pereyra that, necessarily, spatio-temporally related objects are concrete? It seems not. We might believe in *impure* abstract objects which are spatio-temporally located, such as the edge of the table, or the equator. However, this is really no help to the Lewisian realist; for these abstract objects (if there are any) are such that they are ontologically dependent on the existence of concrete objects. The edge of the table may well be abstract; but it only exists if the table exists, and the table is concrete. The equator may well be abstract; but it only exists if the Earth exists, and the Earth is concrete. If there were no concrete objects then

[5] [129, p3]
[6] [ibid.]

there would, it seems, be no abstract objects which are spatio-temporally located. So if there were no concrete objects then there would be *nothing* that is spatio-temporally located. But the Lewisian realist cannot admit the possibility that there is nothing which is spatio-temporally located; hence the Lewisian realist cannot admit the possibility that there could be no concrete objects.

Now Lewisian realism could be tweaked to allow for this possibility. Phillip Bricker[7] suggests that in order to deal with the possibility of island universes — worlds which contain disconnected regions of spacetime — the Lewisian realist should treat the quantifier in the analysis of 'possible' as a plural quantifier. So 'it is possible that p' becomes 'there is some world, or worlds, such that p'.[8] This opens the possibility, says Bricker, of the quantifier ranging over the 'null plurality' of worlds.[9] Possibility then becomes truth at some world, some worlds, or at nothing. In that case Lewisian realism, so altered, allows for the possibility that nothing exists; a fortiori for the possibility that nothing concrete exists. But this is not a very desirable fix: plural quantification over nothing is not too appealing.

However, I don't think there is a real problem here for the Lewisian, because I am not at all convinced that the possibility of there being no concrete objects is something any adequate account of modality has to allow for. The only argument for such a possibility in the literature appears to be Baldwin's subtraction argument. This runs as follows.[10]

1. It is possible that there be only finitely many concrete things.
2. For every concrete thing, it is possible that it not exist.
3. The non-existence of any particular concrete thing does not necessitate the existence of any other concrete thing.

The reasoning that is meant to take us from these premisses to the conclusion that there could have been no concrete objects goes as follows. Consider a world W^1 (possible according to premise 1) where there are finitely many concrete things. Consider one of those things, call it

[7] [22]
[8] [ibid. p44]
[9] [ibid. p47]
[10] [13]

A_1. A_1 might not have existed, and since its non-existence does not necessitate the existence of any other concrete thing we can subtract it from the domain of the world; that is, we can infer that there is a world W^2, possible with respect to W^1, in which everything that was in W^1 exists apart from A_1, and nothing exists which did not exist in W^1. Now consider one of the concrete things in W^2; it might not have existed so there is a world (possible with respect to W^2) in which it does not exist. And so it goes; we reiterate the process until we are left with a world W^n with exactly one concrete thing. Since that one concrete thing might not have existed we infer the existence of a world (possible with respect to W^n) in which there are no concrete things. We then make appeal to the characteristic axiom of S4 — $\Box p \rightarrow \Box\Box p$ — which guarantees that worlds which are possibly possible are possible, and hence that the world with no concrete beings is possible. Is this argument any good? Let's look at its premisses.

5.1.1 Premise 1

If premise 1 is to be true, it must be the case, seemingly, that impure sets with concrete objects in their transitive closure are not concrete. For suppose otherwise, and suppose there is one concrete thing, A. Then there must also be singleton A. If {A} is itself concrete then we now we have two concrete things, A and the singleton of A. But there is also {{A}} (the singleton of the singleton of A); so now we have three concrete beings; and so on *ad infinitum* until we have an infinite number. So it seems impossible to have a finite number of concrete beings, if impure sets are concrete, for whenever we have one concrete thing we have the infinite number of concrete singleton sets that result from it.

For this reason Baldwin is concerned to show that objects such as the singletons of concrete objects are not themselves concrete. Baldwin says that failure to satisfy the identity of indiscernibles is the mark of concreteness; where the identity of indiscernibles is to be taken in its strongest form: if a and b share all their intrinsic properties then a=b. So Baldwin thinks that while for any concrete object there can be a distinct intrinsic duplicate of that object, any duplicate of an abstract object is identical to that abstract object. Since, according to Baldwin, sets, including impure sets, satisfy the identity of indiscernibles in this form, even if their members or ur-elements are concrete, they themselves are not concrete. Baldwin thinks that unit sets

satisfy the identity of indiscernibles, and hence are not concrete, because "the identity of the member of a unit set is an intrinsic property of the set which also determines its identity. Even though there can be two exactly similar (in all intrinsic respects) objects, x_1 and x_2, the unit sets $\{x_1\}$ and $\{x_2\}$ are not in the same way exactly similar since they have different intrinsic properties." Baldwin's point seems to be this: $\{x_1\}$ and $\{x_2\}$ do not share the same intrinsic properties because $\{x_1\}$ has the intrinsic property of having x_1 as a member, and $\{x_2\}$ does not.

I do not find this convincing. I think the property of being a member of $\{x_1\}$ has just as much claim to be an intrinsic property as the property of having x_1 as a member does.[11] If that is right then, if Baldwin's characterisation of concreteness is correct, then either there are no concrete objects or the impure sets with concrete objects in their transitive closure are themselves concrete. For suppose being a member of $\{x_1\}$ and having x_1 as a member are both intrinsic properties: then there could not exist two distinct intrinsic duplicates x_1 and x_2, since x_1 but not x_2 would have the intrinsic property of being a member of $\{x_1\}$, and so, by Baldwin's criterion, there could not exist any concrete objects. Suppose on the other hand that the two properties are extrinsic. In that case $\{x_1\}$ and $\{x_2\}$ fail to satisfy the identity of indiscernibles, since they differ only in the external properties of what their members are; and so, by Baldwin's criterion, they are concrete. So either failure to satisfy the identity of indiscernibles is not what characterises something as concrete, or all entities are concrete and it is impossible that there is a (non-zero) finite number of concrete entities, for given one concrete object C we would also have the infinitely many concrete impure sets with C in their transitive closure.

But I don't think we need let this worry us. Intuitively it does not matter to the thought behind the subtraction argument if impure sets are concrete, because impure sets are ontologically dependent on the concrete objects in their transitive closure: if we get rid of the latter, then, we get rid of the sets as well. It seems irrelevant if there are an infinite number of concrete sets whose existence depends on some concrete non-sets: provided we can subtract the non-sets we will thereby subtract the dependent impure sets as well. Let us change premise 1, then, to 'It is possible that there be finitely many concrete things each of which is not ontologically dependent on some other concrete thing'. The subtraction argument, if sound, would then yield the result

[11] C.f. [127, p162]

that there is a world in which there are no concrete things the existence of which does not ontologically depend on any concrete thing. Such a world, one might think, must be a world in which there are no concrete things at all. For suppose otherwise; then there would be a concrete thing which is ontologically dependent on some other concrete thing. Is that other concrete thing ontologically dependent on another concrete thing? If no, then we've got a contradiction, because the world was said to contain no such things. If yes then is the thing it depends on such that it depends on another concrete thing? And so on. We must stop somewhere; that is, at some point we must answer no to one of these questions, and we'll arrive at a contradiction. So a world in which there are no concrete things which do not ontologically depend on other contingent things is a world in which there are no concrete things simpliciter.

Unfortunately things aren't that simple, because it's not clear that we *do* have to answer 'no' to one of the questions of the form 'does this concrete thing ontologically depend on some other concrete thing?' Why can't we simply keep answering 'yes' ad infinitum? It is an open possibility that there could be an infinite sequence of concrete objects each one of which ontologically depends on the one before it and such that there is no first one; that is, no object whose existence does not depend on any other object. If that is a possibility then the move from 'it is possible that there be no concrete objects whose existence does not depend on any other concrete object' to 'it is possible that there be no concrete objects' fails. One could not argue that once you get rid of all the concrete objects which are not ontologically dependent on others you thereby get rid of all the concrete objects, because there could still be an infinite chain of concrete objects all of which are ontologically dependent on some other concrete object. Instead we want premise 1 to read: 'It is possible that there be finitely many concrete things each of which is not ontologically dependent on some other concrete thing *and* that there be no infinite chain of concrete things such that there is no member which does not depend for its existence on some other concrete thing.'

There is another potential problem with the first premise, and that is that once you have one spatially extended object you thereby, seemingly, have an infinity, for you have all the spatial parts of that object. As Rodriguez-Pereyra says "every . . . object x that occupies a space-time region has infinitely many parts, each of them occupying some of the infinitely many regions

included in the region x occupies. Thus, if there is one spatiotemporally extended . . . object, there are infinitely many."[12] If that is right then how can we avoid there being an infinity of concrete objects; once we have one we have infinitely many.

One might think that this is no problem at all, since it is only *extended* objects that have infinitely many parts; in a world in which there are are no extended concrete objects, merely concrete simples and their sums, there will only be finitely many concrete things. That is true; but I think it would be a bad idea to appeal to this to justify premise 1, because I do not think that the possibility of there being finitely many concrete simples floating in the void is on any stronger an epistemic footing than the possibility of there being no concrete objects at all, in which case an argument from this possibility to metaphysical nihilism will not be suasive.

Rodriguez-Pereyra's own solution to his problem is that when we perform the subtraction and get rid of a thing, we thereby get rid of all its parts. So just as there is no problem with the infinity of singletons which depend for their existence on A, neither is there a problem with the infinity of parts which compose A; when we subtract A we also subtract all of its parts and all the things that depend for their existence of A, or indeed on any of A's parts. But this line of thought won't work; at least not without adding new premisses to the subtraction argument. It does not follow from an objects being concrete that all its parts are concrete. So for the subtraction argument to work the premise would need to be added to the effect that for every concrete object, that object can be *wholly* subtracted, i.e. removed with all its parts. But I do not think that is any good.

The problem is that if we add this premise then the argument is no longer a *subtraction* argument. The notion of subtraction will no longer be doing any work. For if we can get rid of any concrete object and thereby get rid of all its parts, then why not simply, in one step, get rid of that concrete object which is the mereological sum of all the concrete objects.[13] It is itself concrete, presumably; for how could a collection of concrete objects compose an abstract object? Well, if it is concrete, and if concrete objects can be subtracted along with their parts, then every

[12][ibid. p163]

[13]Worries about whether the sum exists are no cause for concern. The argument here does not rely on the truth of unrestricted composition; merely on the possibility of a world which is extensionally equivalent to one in which unrestricted composition is true. Only one who thinks compositional nihilism is necessary will reject the possibility of such a world. But the subtraction argument is not meant to rely on the necessity of compositional nihilism.

concrete object can be removed in one go. Given premise 3 that will get us to the empty world immediately; there is no need to make appeal to the notion of subtraction. So I think to answer Rodriguez-Pereyra's problem in this way is no good; the notion of subtraction is essential to the subtraction argument, so this line of defence would not be defending the subtraction argument but replacing it with a new argument. And the new argument would be no good. The premise that one can remove a concrete object and all its parts is unacceptable if it gets us to the world devoid of concrete beings so easily. Note that the complaint is not that the premise is *false*, but that it is too close to the conclusion to be dialectically effective. The resulting argument basically amounts to 'there could have been no concrete objects, because all the concreta might not have existed, and there be no other concreta'. But that fairly obviously begs the question.

So I think there is a big problem with premise 1. For the subtraction argument not to be question-begging it looks as though we must rely on the possibility of there being finitely many concrete simples. But this certainly doesn't look like a pre-theoretic possibility; this looks like a situation whose possibility or impossibility will be decided by our best metaphysics of modality. In that case it is not a constraint on the best theory of modality that it recognises such a situation as possible; in which case it is not a constraint on the best theory of modality that the premisses of the subtraction argument come out true according to it. And so we have no reason to think that it is a constraint on the best theory of modality that it recognises the possibility of there being no concrete objects.

There is another possible line of defence of premise 1 that is put forward by David Efird and Tom Stoneham.[14] Their solution lies in their definition of concreteness. They take an object to be concrete iff

1) It is spatio-temporally located.

2) It has some intrinsic quality.

3) It has a natural boundary.[15]

[14][37]

[15]The notion of a natural boundary is borrowed from Ted Sider [136]. This third condition should perhaps be changed to 'It has no unnatural boundary'. For consider a world which contains a hunk of homogenous matter unbounded in all dimensions. It has no natural boundary, because it has no boundary: therefore it is not a concrete object according to Efird and Stoneham. But I should not count as a metaphysical nihilist simply in virtue of believing that this situation is possible. We should count this object as concrete, then; and one way to do this is to demand not that concrete objects have natural boundaries, but to demand that any boundary they have is natural.

Consider a white sheet of paper with a red circle on it. Both the paper and the circle have a natural boundary; but the red circle whose center-point is the same as that of the circle just mentioned, but whose radius is, say, half of that first circle's radius, does not have a natural boundary: the boundary of this inner circle does not contrast with its surroundings in the appropriate way. Now consider a world in which the only thing that exists is that white sheet of paper with the red circle on it. It is extended in space, and so it has infinitely many proper parts. But there are not infinitely many concrete objects in this world, given Efird and Stoneham's definition of concrete; there are only two concrete objects in this world, for there are only two objects which have natural boundaries: the paper, and the part of the paper which the circle is drawn on. The other parts of the paper will not count as concrete, and hence not fall under the quantifiers in (A1). My problem is that if *this* is what it is to be concrete, then I think they introduce extra problems for premise 3 (problems over and above those that we will see below).

Efird and Stoneham actually formulate their version of the subtraction argument using only two premisses. They basically lump premisses 2 and 3 above together, which gives them (B) below.

(B): $\forall w_1 \forall x(E!xw_1 \rightarrow \exists w_2(\neg E!xw_2 \wedge \forall y(E!yw_2 \rightarrow E!yw_1)))$

Where 'E!aw' is to be read as 'a exists at world w', w_1 and w_2 range over possible worlds, and x and y range over possible concrete objects.[16]

(B), together with premise 1 — that there could have been finitely many concrete objects[17], seemingly entails their conclusion $\exists w \forall x \neg E!xw$, which says that there is a world in which none of the possible concrete objects exist; which, of course, is a world in which no concrete objects exist.[18] Below I will argue that principles like (B) are not supported by the intuitive principle they appear at first to be supported by.[19] But even putting that worry aside, I think that we should not accept Efird and Stoneham's argument. The problem is that (B) does not look

[16] I.e those objects among the possibilia that are concrete.

[17] The particular premise Efird and Stoneham use is that there could have been two concrete objects: $\exists w \exists x \exists y((E!xw \wedge E!yw) \wedge \forall z(E!zw \rightarrow (z = x \vee z = y)))$. But their argument is valid for any natural number n in the place of 2, if it is valid in this case.

[18] For if a concrete object existed at this world then it would not be a possible object, in which case the world would not be a possible world; but that is ruled out because w only ranges over possible worlds.

[19] The relationship of (B) to the principle (2), which Efird and Stoneham appeal to in justification of (B) as we will see below, is like the relationship between what I call ultra-strong DNC and strong DNC below. In each case the latter principle is quite intuitive, but does not justify the former principle.

remotely plausible given the definition of concreteness that Efird and Stoneham use to justify premise 1.[20] In that case, if premisses 1 and (B) are both to be accepted then the objects which the non-world variables in their formulation of premise 1 range over are not the same objects that the non-world variables in (B) range over; and if that is true then their argument is invalid.

I will now argue that (B) is false when the quantifiers range only over those things which have natural boundaries. This point is perhaps best seen by considering the rather strange tribe of people known as the Qube. The Qube worship their god, Qubec, which is a large cube made of homogeneous matter. The Qube believe that Qubec has the size it has essentially: in particular, they believe that anything smaller than Qubec would not be Qubec. Part of the reason for this is that Qubec contains within him an infinite number of other gods. So suppose you chip off some of the matter from the surface of Qubec, but in such a way as you are left at the end with a cube of less volume. This cube is not Qubec, but rather one of the other gods who is contained within Qubec — call it Qubec*. Qubec* has its size essentially as well; were we again to chip off some matter from its exterior we would destroy it and be left with yet another god — call it Qubec**. And the Qube believe this process can go on indefinitely: you can keep getting rid of a god from the world without ever getting rid of all the gods. Since the biggest god is a concrete object, this means that the Qube believe you can keep on removing a concrete object from the world without getting to a world devoid of concrete objects. So far so consistent with metaphysical nihilism; because Efird and Stoneham do not need to deny that there are *some* processes of subtraction whereby you will not get to an empty world, all they require is that there be some process that does. But unfortunately, while the Qube believe that any one of the gods might not have existed[21], they think it is necessary that there are *some* gods. In that case, the Qube are committed to the denial of metaphysical nihilism. They think that in every possible world there is a concrete object; for one of the gods always has a natural boundary, and since all the gods both have intrinsic qualities and are spatio-temporally located, one of the gods will be a concrete object.[22] And so the Qube had better deny either premise 1 or (B). Clearly it is (B) that the Qube deny. The Qube think that removing a concrete object can necessitate the introduction of some new concrete object. But are the Qube, in thinking this, denying some

[20] Whereas on an ordinary definition of concreteness the problem with (B) is subtler.
[21] And hence they accept the claim that, necessarily, all concrete objects are contingent existents.
[22] Which one is concrete will vary from world to world obviously.

pre-theoretic modal intuitions? I don't think so. Efird and Stoneham locate the justification for (B) in the following principle which they call (2).

(2): Necessarily, if there are some concrete objects, there could have been fewer of those concrete objects (and no other concrete objects).

Now in fact, as will become clear in my discussion of premise 3 below, I don't think principles like (2) do serve to justify principles like (B). But the point there is rather subtle; for now let me make the less subtle point that (2) *definitely* doesn't justify (B) when (B) is taken to be about concreta in the Efird and Stoneham sense.

(2) is very plausible when we do not demand that concreta have natural boundaries. In that case it looks as though whether or not a thing is concrete is essential to it, in which case the only way (2) can fail is if the non-existence of some concrete thing necessitates the existence of *some* concrete thing[23] which doesn't in fact exist. So it seems to me that what is bad is if the non-existence of some concrete thing necessitates the existence of some new concrete thing in the sense that it necessitates the existence of some concrete thing that is a *new existent*: something that didn't exist in the world we started from. It does not seem at all bad to me if the non-existence of some concrete thing necessitates the existence of some new concrete thing in the sense that it necessitates that something is *newly concrete*: i.e. concrete in the new world but not concrete in the world we started from. At least, that is not bad if concreteness is the extrinsic property that Efird and Stoneham think it is. So the Qube do not seem to me to be denying (2) in the sense in which (2) is a pre-theoretical modal intuition. They deny (2) only in the sense in which it should be denied: they deny that subtracting something from a world cannot change how the rest of the things in that world are with respect to a certain extrinsic property — concreteness as understood by Efird and Stoneham. They do not deny the very strong intuitive thought behind (2), for they do not think the non-existence of some thing can necessitate the existence of some thing which does not in fact exist. I conclude that Efird and Stoneham's subtraction argument is unsuccessful. They solve the problem for premise 1 by appealing to a notion of concreteness which makes being concrete an extrinsic property; but with concreteness thus understood (B) appears to be false, and not motivated by (2). (B) *might* be thought to

[23]Not some particular concrete thing: just some concrete thing or other.

be true if we take concreteness to be an intrinsic property that all the infinitely many parts of concrete things have, and which as a matter of necessity is had essentially or lacked essentially (although see the discussion of premise 3 below). But when we understand concreteness this way the possibility of there being finitely many concrete objects is not immediately obvious, and is certainly not more obvious than the possibility of there being no concrete objects; and so the argument with premise 1 construed thus is not going to be suasive. (B) and premise 1 can only both be secured, then, if the quantifiers in (B) range over all and only the things which are concrete in the intrinsic and essential sense, and the quantifiers in premise 1 range over all and only the things that are concrete in the extrinsic and accidental sense. In that case Efird and Stoneham's argument will be invalid, because they make the illegitimate move of instantiating a quantifier in premise 1 and then using that constant to instantiate a quantifier in (B). Since the quantifiers do not have the same domain, this step is invalid.

5.1.2 Premise 2

What of premise 2? Baldwin thinks it is obvious that each of the concrete objects might not have existed. I am not so sure; one might think, for example, that space is a necessarily existing concrete object. We should note at this stage, however, that premise 2 as it is stated with regards to concrete objects is not enough for the subtraction argument to go through without strengthening our logic from S4 to S5. The three premisses of the argument with respect to concrete objects leave it open that it is possible that (at least some of) the concrete objects might have existed necessarily (but still be concrete). If that were the case, then even though each of the concrete objects is such that it might be removed, we have no assurance that in the worlds in which one of those objects is removed, each of the remaining objects can still, *in that new world*, be removed; and so there is no guarantee that we can keep reiterating the process of subtracting concrete objects to get to a world void of them. In order to make the subtraction argument as it concerns concrete objects sound we have two options: either we can strengthen premise 2 to read '*Necessarily* each concrete object might not exist', or alternatively we can leave the premisses as they are and make appeal to the characteristic axiom of S5: $\Diamond p \rightarrow \Box\Diamond p$.[24]

[24]That some concrete object is contingent but might have been necessary is of course compatible with the B axiom: $p \rightarrow \Box\Diamond p$. B implies that if it is possible that there is a necessary being, a, then a actually exists; but it

That ensures that objects which can be removed at one world in this chain can still be removed at worlds further on; since if it is possible that they do not exist then, by the S5 axiom, it is necessarily possible that they do not exist. (Accepting premise 2 as necessary is consistent, of course, with denying the S5 principle; for one might well think that it is necessary that every actual concrete could not exist but still think that there could have been some concrete object which does not actually exist and which would have been a necessary existent.)

Baldwin holds that premise 2 is necessary, thus he resists the pressure to adopt S5. Why does he think this? Well remember that Baldwin thinks that the mark of concreteness is failure to satisfy the identity of indiscernibles in its strongest form. In that case, he says, the identity of an object is not determined by its intrinsic properties, even though the kind of object it is is so determined. So that an object a is concrete is determined by its intrinsic properties, but that it is the very thing a is not so determined. But for any necessary existent a, thinks Baldwin, the fact that a is necessary follows from what kind of thing it is. Thus the fact that a thing is necessary is determined by what kind of thing it is. Now suppose that there were a necessary concrete existent, call it c. The fact that c is necessary is determined by its intrinsic properties, but the fact that it is c is not. But Baldwin thinks this is absurd; for necessary existents the properties which determine an objects existence must also, he thinks, determine its identity. So the properties which determine that c exists necessarily must also determine that c is c. But if c is concrete then they do not. It follows that there can be no concrete necessary existents.

What should we make of this argument? Well we have already seen reason to doubt Baldwin's mark of concreteness; but even granting that I think the argument can be resisted. Baldwin says it is uncontentious that for a necessary existent the kind of thing that it is determines that it is necessary. Moreover, he needs that to be necessarily true. But that sounds like exactly the principle that I am going to reject if I think that there could be concrete necessary existents. If I think that, why couldn't there be objects of the same kind some of which are contingent and some of which are necessary (even if I don't believe that any actual contingent object could be necessary)? Also dubious is the argument for the claim that the properties which determine the identity of a necessary existent must also determine that it is necessary. Again, Baldwin needs

---

does not follow that a is actually necessary.

this premise to be necessarily true, but has given us no argument that it is even contingently true. So I don't think Baldwin's argument for the necessity of premise 2 is convincing. A better option for those who wish to show that there could be no concrete objects, I think, would be to make appeal to the E axiom — $\Diamond p \rightarrow \Box\Diamond p$ — and accept as contingent premise 2 (or at least remain agnostic as to its modal status).

5.1.3 Premise 3

What of premise 3? I think it should be rejected, and that it sounds plausible only when confused with a weaker premise that does not suffice for the validity of the subtraction argument. What is extremely plausible is the following principle: that the non-existence of any particular concrete thing does not necessitate the existence of any other *particular* concrete thing. That is, for any particular concrete thing A, it is not the case that in all worlds in which A does not exist some other concrete thing B exists. So what is ruled out is that there is one concrete thing which exists in all worlds in which A does not exist. But that principle is compatible with *some* concrete thing existing in all the worlds in which A does not exist; provided that it is not the same thing in each of those worlds.

It might be the case, for all the principle says, that there is no necessary connection between the existence of A and the existence of B, and likewise that there is no necessary connection between the existence of A and the existence of C (a concrete being), but that necessarily if A exists, either B or C exists (and, of course, perhaps both). Premise 3 needs to be strong enough to rule this situation out for the subtraction argument to work. But it is the weaker principle, I think, that has a right to be called the denial of necessary connections between distinct existences. So attraction to this Humean/Lewisian principle does not support the stronger thesis. It is easy to read premise 3 as amounting to the weaker, relatively uncontroversial principle; but that will not get the desired conclusion. Premise 3 has to say not only that the non-existence of each of the contingent things does not necessitate the existence of any other particular thing, but that it does not necessitate the existence of anything *per se*.

So there are two principles on the table so far: what I will call the weak and medium denial of necessary connections:

Weak DNC: For each thing a, the non-existence of a does not necessitate the existence of any other *particular* thing.

Medium DNC: For each thing a, the non-existence of a does not necessitate the existence of any other thing *per se*.

But a debate between Alexander Paseau[25] and Gonzalo Rodriguez-Perayra[26] shows that even medium DNC is ambiguous. It could mean that each thing in a world w is such that its non-existence does not entail the existence of even one of the things in w, or it could mean that each thing in a world w is such that its non-existence does not entail the existence of any other possible thing, whether in w or not. Paseau shows that the subtraction argument is invalid on the former reading but doesn't consider the latter, which is Rodriguez-Pereyra's intended reading. Rodriguez-Pereyra says premise 3 is to be read as follows, which I will call strong DNC.[27]

Strong DNC: "The non-existence of any of the x_i that exist in w_1 does not necessitate the existence of any other concrete object, whether or not these exist in w_1."[28]

He follows this with what is supposed to be a possible worlds translation of the same principle. I will call this ultra-strong DNC, for reasons to become clear.

Ultra-strong DNC: "[F]or all worlds w and for all the concreta in x_i in w, if x_i exists in w then if there is a world w^* where x_i does not exist, then there is a world w^{**} where the only existing concreta are those of w except x_i."[29,30]

The problem is, Rodriguez-Pereyra's possible worlds translation is stronger than the English principle he gives: strong DNC is weaker than ultra-strong DNC. He is correct that the subtraction argument is valid given the possible worlds principle, ultra-strong DNC; but if the only plausibility this principle gets is from confusion with the English principle, strong DNC, as I

[25] [99]
[26] [128]
[27] I have changed each occurrence of 'concrete*' with 'concrete' throughout. The star is not needed as a result of the arguments concerning premise (1) above.
[28] [ibid. p172]
[29] [ibid.]
[30] Ultra-strong DNC is very similar to Efird and Stoneham's principle (B), which we saw in the above discussion on premise 1, and the points I make against ultra-strong DNC also apply to (B) *even* when the quantifiers in (B) are taken to range over things which are concrete in the sense in which concreteness is an intrinsic property that is had or lacked by things essentially.

suspect is the case, then we should not accept the argument as sound.

Why does strong DNC not validate the subtraction argument? It might be the case that for any object, you can remove that object without any new object being introduced, but it still be the case that there is some crucial minimum number of objects that there can be. Suppose, for example, that there is a brute necessity that there cannot be less than three concrete objects. That ought to be in conflict with the premisses of the subtraction argument. But it is not, if premise 3 is strong DNC. Consider a world W_1 in which the only things that exist are A, B and C. Suppose I think the following worlds are possible relative to W_1:[31]

W_2, in which the only things that exist are A, B and D
W_3, in which the only things that exist are A, B and E
W_4, in which the only things that exist are B, C and D
W_5, in which the only things that exist are B, C and E
W_6, in which the only things that exist are A, C and D
W_7, in which the only things that exist are A, C and E

I stipulate that there are no worlds possible relative to W_1 containing less than three objects. In that case it is impossible, at W_1, that there be less that three objects. In which case the premisses of the subtraction argument had better not be true at W_1. Well premise 1 is certainly true, and we can grant that 2 is true. So strong DNC must be false at W_1. But it needn't be false. It can be true at W_1 that for each thing A, B and C, the non-existence of each of those things does not entail the existence of any other thing. For there are worlds which are possible relative to W_1 in which A, B and C exist and some other things exist. In particular, there is a world W_8, possible relative to W_1, in which the only things that exist are A, B, C and F. Consider now the world W_9, possible relative to W_8, in which the only things that exist are B, C and F. The only difference between W_8 and W_9 is that A doesn't exist in W_9. Nothing exists in W_9 that does not exist in W_8. In that case it is true at W_1 that the non-existence of A does not entail the existence of anything, since we have shown that it is possible at W_1 for A not to

[31] All the things here are concreta.

exist and no other thing to be introduced.[32] Exactly the same story can be told about B and C; so it can be true at W_1 that for each of the things that exist their non-existence does not entail the existence of anything else; i.e. for each of those things it is possible that it not exist and that no new thing exist. So premisses 1 and 2 of the subtraction argument can be true at a world in which strong DNC is true and yet it be true that the minimum number of concrete things that exist is 3. In which case strong DNC is not strong enough for the subtraction argument to be valid.

So premise 3 needs to be even stronger than strong DNC. How should it be strengthened? What went wrong with strong DNC is that to show that the non-existence of a thing a in a world w does not entail the existence of any new thing we need only show that there is a possible world in which a could fail to exist and no new thing be introduced. But, crucially, that possible world needn't be w. The defender of the subtraction argument must rule out such a move; but they will not do so by talking about what the non-existence of an object fails to entail.

Let no one think the following is a good response to this problem: that the thought that premise 3 is aiming to capture is that an object can be removed *in that world* without introducing any new objects, and that is what is being broken in the above example. That is not being broken. To show that an object a *can* be removed in w without introducing any new objects we need only show that it is *possible in w* that a be removed without introducing any new objects. Which is exactly what we did in the above example: we described a world possible relative to w in which one can remove a without introducing any new objects.

The thought the defender of the subtraction argument is trying to capture, I take it, is that each object can be removed from a world, and everything else about the world remain the same. That is, given a world w, any x in w can be removed and there be no other difference. That is the thought that is being violated above. The above situation describes a situation in which the non-existence of any of the objects in W_1 does not entail the existence of any other things; but it does so only by showing that the domain of W_1 could be bigger, and that were it bigger each of

[32] One might object that all we have shown is that it is possibly possible, since W_9 is possible relative to W_8, which is possible relative to W_1. But that would not be a good objection for two reasons: (1) even if one denies S4, it seems that we need to have good reason to say that some situation is possibly possible but impossible, and no reason is forthcoming here, and (2) certainly this is no response the defender of the subtraction argument can make, since the argument assumes S4.

those could fail to exist without some other thing existing. This, the defender of the subtraction argument might think, is cheating. It is no good that each of those of objects could fail to exist *were the domain of the world bigger* and no other object introduced. What the defender of the subtraction argument wants, and what the above situation rules out, is that each of the objects in a world could fail to exist, and *the world be as it is in all other respects*, and no other object be introduced.

It is ultra-strong DNC, then, that is not true at W_1 above, for it says that given a world w with a domain D, for every concrete x in D, there is a world possible relative to w whose domain is $D - \{x\}$. Ultra-strong DNC is false at W_1, and is needed for the validity of the subtraction argument. But what reason is there to hold ultra-strong DNC? It is not, as we have seen, a legitimate possible worlds paraphrase of the principle Rodriguez-Pereyra takes his argument to rely on; and it no longer bears any relation to anything that might legitimately be called the denial of necessary connections between distinct existences. The bullets that I would be forced to bite in denying weak or even strong DNC I am not forced to bite in denying ultra-strong DNC. If I deny weak or strong DNC then I commit myself to thinking that particular objects have, in some sense, the power to replace themselves; that their failing to exist can in some sense be responsible for the existence of some other thing. That is certainly a bullet to bite, although it's not clear how bad a bullet. But no such commitment is brought in denying ultra-strong DNC. Ultra-strong DNC, unlike weak and strong DNC, is not a principle about what would have to have happened if some particular thing had failed to exist. It is not, as the other two principles are, in any real sense a principle about the objects in a world, and the connections that they have or fail to have with other (perhaps merely possible) objects; rather it is a principle about the domain of the world as a whole. And I suspect the only credibility it gets is from confusion with either weak or strong DNC

Weak and strong DNC are principles about the essences of objects. In particular, they are denials of certain essentialist theses. Weak DNC denies that there is any pair of possible objects such that it is in the essence of one that it exists if the other fails to exist. Strong DNC denies that there is any object which is essentially such that if it does not exist there must be some new thing that exists. But ultra-strong DNC is a different sort of principle altogether. It is

a principle concerning the domains of possible worlds, not the objects in those domains. Now maybe the defenders of the subtraction argument have some strong intuitions not just about the essences of objects but of the essences of domains, but they have yet to sufficiently argue their case.

I conclude that the subtraction argument fails to establish that metaphysical nihilism is entailed by pre-theoretic modal intuitions. Premise 2 is acceptable, at least given S5. But the acceptability of premise 1 relies on the assumption that there could have been only finitely many concrete simples; and that is a situation whose possibility or impossibility is to be settled by the best account of modality: its possibility is not pre-theoretical. And I think we have no reason to accept premise 3 in the version that is needed to make the argument valid. This is a premise that talks about the domains of possible worlds, and as such has no claim at all to be a pre-theoretical modal claim.

Given the failure of the subtraction argument, we have no reason to think that any adequate metaphysics of modality must be compatible with the possibility of there being no concrete objects. In particular, then, it is no objection to Lewisian realism if it rules out metaphysical nihilism.

5.2 Circularity in Analysis

The second objection I will look at is that Lewisian realism is not successful because it is not an analysis. The objection is that the analysis only succeeds if there is modal vocabulary in the reductive basis, and hence it would not be a genuine reduction. This objection originates with Bill Lycan. Lycan[33] poses the following problem for a Lewisian realist. One claim of Lewisian realism is the following equivalence scheme PW which says that a propositions being possible is equivalent to it being true at some world (where 'Wx' is read as 'x is a world' and 'T⌜p⌝x' is read as 'p is true at x').

PW: $\Diamond p \leftrightarrow \exists w(T\ulcorner p \urcorner w)$

Now, says Lycan, for Lewisian realism to be a successful analysis of modality it must rule out

[33] [86]

impossible worlds from the domain of quantification — otherwise the theory would incorrectly claim that impossibilities are possible. For example, consider a debate between a Lewisian realist and a Meinongian. The Meinongian believes that there are such things as round square cupolas; she does not claim that they are possible, but that does not mean that there are none for the Meinongian: they do not exist but they subsist as mere impossibilia. Now the Lewisian realist, agreeing with the Meinongian that round square cupolas are impossible, has to reject the claim that there are, anywhere in logical space, round square cupolas. At no world is there a round square cupola, thinks the realist. However, says Lycan, herein lies the primitive modality. What is to stop there being a world in which there are round square cupolas other than that the worlds we are quantifying over in PW are possible worlds and not impossible ones? A related worry is considered by Scott Shalkowski.[34]. Not only must the worlds we quantify over in PW not include any impossible worlds, they must be complete in the sense that our domain of quantification in PW must include every possible world. The Lycan-esque conclusion is that this is only the case if 'world' is understood as the primitively modal 'possible world'. For otherwise what would ensure even that there is more than one world, in which case PW would give us the false conclusion that everything is necessary?

In both cases the problem is this. For Lewisian to be successful PW has to be materially adequate. That is, Lewisian realism must entail that among the worlds quantified over in PW there is one corresponding to every possibility and there are none corresponding to an impossibility. But Lewisian realism can only do this, the objection goes, if PW is understood as quantifying over possible worlds. Hence Lewisian realism can only be successful if it relies on a primitive modality — namely, the notion of a possible world. Lycan concludes, "For [Lewis] 'world' just means 'possible world'. Thus it is itself a modal primitive even though it is not spelled like one."[35]

Lewis actually goes some way to giving ground to this objection, for he says of the worlds that PW quantifies over that "every way the world could be is a way some world is."[36] If this is intended as a *definition* of what a world is then it must fall victim to Lycan's objection, for 'way

[34][133] Also c.f. Divers [30]
[35][ibid. p224]
[36][66, p86]

the world could be' is clearly modal. But, as Richard Miller points out[37], Lewis has the material to provide us with a definition of world which does not seem to use any modal terms whatsoever. First of all we take Lewis' definition of what it is for any individuals a,b to be worldmates: a and b are worldmates iff a and b are spatiotemporally related. Then we simply define 'world' to be a maximal mereological sum of worldmates. So defined, the right hand side of PW quantifies over mereological sums, and mereological sums are not, it seems, primitively modal.

But, thinks Lycan, such a move is just shifting the problem. Grant that there is nothing intrinsically modal about the notion of a mereological sum. The problem, thinks Lycan, is that in specifying what we are taking mereological sums of we are appealing to the notions 'individual' and 'spatiotemporal relation'. And, says Lycan, now the problem arises that for Lewisian realism to be materially adequate these must be understood as the primitively modal 'possible individual' and 'possible spatiotemporal relation'. He says that "nothing but the impossibility of round square cupolas in the first place keeps a round square cupola from being spatiotemporally related to another, perhaps less exotic object."[38] So if the realist wishes to deny the existence of a world in which there are round square cupolas — which she does — then she must claim that round square cupolas are not the types of things that can go into the mereological sums in question. But the only justification she would have to claim this, thinks Lycan, is to understand 'individual' as 'possible individual'. Lycan concludes that "Lewis would have to amend ['a world is a maximal mereological sum of worldmates'] to read, 'A world is a maximal mereological sum of possible worldmates'.[39] This would rule out the existence of worlds containing impossible individuals (and a similar move would rule out worlds with impossible spatiotemporal relations between individuals), but it does so at the price of relying on the clearly primitively modal notion 'possible worldmate', thus defeating the main advantage of Lewisian realism.

I don't think that Lycan's objections work; he is asking for far too much from the realist. Lycan's argument assumes that the Lewisian realist has to rule out impossible worlds, or impossible worldmates given the identification of worlds with maximal mereological sums of spatiotemporally related individuals. But in fact the realist has to do no such thing. The scheme PW says

[37][91, p477]
[38]Lycan op. cit p 212.
[39]ibid. p212-213.

that any world which exists is a possible world, any individuals that exist (at any world) are possible individuals, and any spatiotemporal relation between any two individuals is a possible spatiotemporal relation. But the direction of explanation here is not from 'such and such is possible' to 'there is a world containing such and such' — that would indeed require 'world' to be understood as primitively modal. Rather, the direction of explanation is from 'such and such is a world' to 'such and such is possible', and this requires no such primitive modality. For any world w, w is a way the world could be, not because 'world' is defined as 'a way the world could be' but because 'a way the world could be' is defined as 'a way some world is'. That's what it is to analyse possibilities in terms of worlds. Lycan challenges Lewis to rule out impossible worlds, and later to rule out the existence at some world of an impossible individual or an impossible spatiotemporal relation between individuals. But this gets the direction of explanation the wrong way round. If there exists a mereological sum of individuals one of which is a round square cupola then, if Lewisian realism is correct, round square cupolas simply *are* possible. If there exists a world at which a contradiction is true then, if Lewisian realism is correct, contradictions *are* possible. So the realist has nothing to rule out. Either p is not true at any world or p is possible; PW only quantifies over existing things, so either way there is no danger of PW quantifying over an impossible world.[40]

As we said above, Lewisian realism must entail that among the worlds quantified over in PW there is one corresponding to every possibility and there are none corresponding to an impossibility. But if Lewisian realism is true then PW is true, and if PW is true then this challenge is met trivially. There is no need for the Lewisian realist to worry about whether or not PW is quantifying over impossible worlds. If you accept Lewisian realism then what it is for a world to be impossible *just is* for it not to be quantified over by PW. Similarly there is

[40]Let me make a brief qualification. Lewis in fact does believe in impossible individuals, but he doesn't believe that there is any world at which there exists an impossible individual; see [66, p210-211]. It is the latter claim that trivially cannot be true under Lewisian realism, and it is that claim that Lycan is challenging Lewis to rule out. The only impossible individuals Lewis believed existed were trans-world individuals: individuals which have a part at one world and another part at another distinct world. Such individuals must exist according to Lewis' ontology, because he claims that for any individuals a and b there is an individual which is the mereological sum of a and b; so if a exists at world w and b exists at world w^* (such that $w \neq w^*$) then there exists a trans-world individual which is the mereological sum of a and b. But such an individual is not a possible individual, since to be a possible individual is to exist at a world, and trans-world individuals do not exist at any world (for by definition, at most they have a proper part at any world). So Lewis does believe in impossible individuals, but not the kind that will pose any problem for Lewisian realism. The existence of trans-world individuals does not give us the bad result that Lycan is challenging Lewis to rule out: that some impossibility is possible.

no need for the modal realist to worry about whether or not the worlds quantified over in PW take care of every possibility. If you accept Lewisian realism then what it is for something to be possible *just is* for it to be true at some world quantified over by PW. Just as we can be sure that we will never discover that the actual world is not a logically/physically etc possible world since logical/physical etc necessity is simply defined as what must be true given the actual laws of logic/physics etc, so can the genuine modal realist be sure that she will never discover that there exists a world which is not possible, since possible world is simply defined as how some world is.

There may be a feeling here that I am doing something illegitimate in my defence of Lewisian realism against Lycan's objection: namely appealing to Lewis' analysis of possibility in terms of truth at a world in order to defend it. Obviously I can justify the material adequacy of an analysis of Φ in terms of Ψ if I am allowed to appeal to the truth of the biconditional $\Phi(x) \leftrightarrow \Psi(x)$. But, the objection might go, our only reason for believing that biconditional is if we believe the analysis, so we cannot use the biconditional to defend the analysis. There is a certain amount of truth in this objection, but as a response to my criticisms of Lycan it is misplaced. Lycan's objections are not aimed at showing the falsity of the scheme PW. Rather, they are aimed at showing that its truth presupposes a modal primitive in the theory Lewisian realism. Lycan's argument is only an argument against the truth of Lewisian realism indirectly; if it can be shown that Lewisian realism does not provide an analysis of modality then that weighs heavily against it, since that is one of its main advantages. The objection, then, relies on the claim that even if Lewisian realism is true, it must contain a modal primitive. But in that case one is allowed to assume that Lewisian realism is true in order to resist Lycan's objection. To show that the conditional 'If GMR is true it contains a modal primitive' is false I am entitled to assume the antecedent and show that, on that assumption, the consequent is false.

5.3 The epistemological and modal irrelevance objections.

This section focuses on two objections to Lewisian realism that I think are successful: the epistemological objection and the modal irrelevance objection. The epistemological objection is

that on the account the Lewisian realist gives of the truth-conditions of modal claims, knowledge concerning modal claims would be unobtainable. The modal irrelevance objection says that the account of the truth-conditions given cannot be correct because what goes on at a Lewisian world is irrelevant to whether or not something is possible — Lewisian worlds, even if there are such things, have nothing to do with modality. I will begin by reviewing the case for each objection, and Lewis' response to the objections; we will start with the modal irrelevance objection.

5.3.1 The modal irrelevance objection

The objection charges that Lewisian worlds have nothing to do with modality; what goes on at some world is simply *irrelevant*, so the charge goes, to what is or is not possible. van Inwagen, for example, says that Lewis must "face the problem of explaining what these things [these maximal mereological sums of spatio-temporally related individuals] would have to do with modality if there were any of them."[41] The objection is put across in more detail by Michael Jubien[42]

> [I]t certainly makes sense to imagine that there *are* entities rather like the world but that lie outside of its spatiotemporal limits. Given this much, it would be very natural, though hardly mandatory, to say that our world could have resembled these entities in various ways. But why should we suppose that for *any* way our world could have been, there is an entity of this special sort that *is* that way? . . . What can they [Lewisian worlds] have to do with possibility?

Note that there are two different lines of objection in this quote from Jubien. There is the question as to what Lewisian worlds have to do with modality (or rather, what they would have to do with modality, were they to exist), and there is the question as to why we should think there is a Lewisian world for every way the actual world might have been. Clearly they are not the same question, and one might wonder why they are lumped together. I hope to show later, however, that although the two objections are distinct, they are related, and are lumped together rightly.

[41] [146, p119 Fn.15]
[42] [57, p305]

Lewis responded to the charge of irrelevance by saying that he *had* indeed explained what Lewisian worlds have to do with modality, namely "by saying that the modal operators are quantifiers over them."[43] Chihara responds that this does not answer the objection, because it presupposes the truth of the analysis that is up for debate. He says[44]

> [Lewis'] response to the Modal Irrelevance Objection was that he had already explained what worlds have to do with modality by saying that the modal operators are quantifiers over them. But that is just what is being disputed . . . So, from the point of view of his disputers, Lewis was simply begging the question against his opponents. . . Lewis responds to [the] objection from the perspective of one for whom the analysis is already beyond question.

In order to know who has won this debate we need to know whether or not Lewis is entitled to respond to the modal irrelevance objection "from the perspective of one for whom the analysis is already beyond question." You might think he obviously is; for who's perspective is he to respond from if not his own? But I think a case can be made that Lewis is *not* entitled to respond from his own perspective. Before I attempt to make this case, however, let us look at the epistemological objection.

5.3.2 The epistemological objection

The epistemological objection to Lewisian realism claims that if Lewisian realism gives a correct analysis of the modal then modal knowledge would be mysterious, since it is a mystery how we could have knowledge of what goes on at non-actual worlds. Since we should not accept an analysis of modality which does not cohere with an adequate epistemology of the modal it follows that we should not accept Lewisian realism. This objection is made by Tom Richards who says that according to Lewisian realism "the truth conditions [of modal claims] are such that, for any given statement, it is impossible in general to determine whether they are met and hence whether the statement is true."[45]

[43] [Lewis op cit. p98]
[44] [26, p95]
[45] [124, p109]

Why can we not know about what happens at non-actual worlds? The general idea, I take it, is that in order to know facts concerning concreta we must have causal interaction with those concreta; that while I might be able to attain a priori knowledge concerning the existence of abstracta, and about the properties of abstract objects, a priori knowledge concerning the existence of concreta, or the properties of concrete objects, is impossible. Knowledge of concreta is all a posteriori, obtainable only via sensory investigation; since we can only have sense data of things which causally interact with us it follows that we cannot have sense data of objects which are causally disconnected from us. Since we are in no causal relationship with any non-actual being it follows that we cannot have knowledge concerning the concrete non-actual beings that Lewis tells us exist. Put thus, the objection relies on the following two claims: that a priori knowledge concerning concreta is impossible, and that a posteriori knowledge of concreta requires causal interaction with those concreta.

Lewis responds thus: "I think it is true that causal acquaintance is required for some sorts of knowledge but not for others. However, the department of knowledge that requires causal acquaintance is not demarcated by its concrete subject matter. It is demarcated by its contingency."[46] His claim is that while knowledge about what concreta *actually* exist, or knowledge concerning the properties of particular concrete individuals, requires causal interaction with the relevant concreta, that is because that is knowledge of *contingent* matters of fact. Knowledge of the non-contingent claim that there is unrestrictedly a talking donkey, on the other hand, requires no causal interaction with any talking donkey.

Note that again Lewis has responded to an objection to Lewisian realism by relying on claims that the non-Lewisian is unlikely to think are true, for even if it is true that causal acquaintance with some subject matter is only a condition for knowledge concerning that subject matter if the subject matter is contingent, this response is not going to convince Lewis' opponents that knowledge about non-actual talking donkeys does not require causal interaction with such things, since it is *only according to the Lewisian realist* that facts concerning the existence of non-actual talking donkeys are necessary.

For the non-Lewisian realist, facts concerning what concrete beings there are — by which

[46] [66, p111] See also [32, p158-161].

I mean what concrete beings there are *unrestrictedly*, not merely what concrete being there *actually* are[47] — are contingent.[48] It is for precisely this reason that the non-Lewisian can draw a motivated distinction between knowledge of abstracta and knowledge of concreta. Why is there a causal constraint on knowledge of the existence of concreta and not on knowledge of the existence of abstracta? Because facts concerning the existence of abstracta are necessary. Given the B axiom — $p \rightarrow \Box\Diamond p$ — it follows that the possibility of a statement concerning the existence of an abstract object is sufficient for the truth of that statement.[49] Since the space of possibilities concerning the existence of abstracta can be known a priori, it follows that we can know a priori what is true of the domain of abstract objects. In contrast, we can perhaps know a priori what is possible concerning the (unrestricted) domain of concrete objects, but have no idea which of those possibilities is realised, and in order to have such knowledge we need to rely on empirical experience. That is why, for the non-Lewisian, there is a distinction between knowledge of concreta and knowledge of abstracta. Lewis, of course, will deny the disanalogy; so long as we are talking about what concrete objects there are *unrestrictedly*, possibility does indeed suffice for truth in the same way as it does for abstracta; for possibility is sufficient for truth at some point in logical space, and the unrestricted domain of concreta includes every concrete being in logical space. All well and good; but that simply shows that the disanalogy between knowledge of concreta and knowledge of abstracta only holds good if one is not a Lewisian realist. It will not convince the non-Lewisian that the disanalogy is no good, and hence to the non-Lewisian knowledge concerning non-actual worlds still looks mysterious.

So we haven't been given a reason independent of Lewisian realism to accept that modal knowledge is not mysterious according to Lewisian realism; we have merely been given a reason to think that *according to Lewisian realism* such knowledge is not mysterious. Lewis acknowledges the point. He says "So we have the desired boundary between knowledge that does and that

[47] For the non-Lewisian of course, this will most likely be taken to be a distinction without a difference. Nevertheless, it is important to emphasise that there is a genuine disagreement with Lewis that is nor merely a disagreement about ontology. There is a disagreement between the Lewisian and the non-Lewisian over whether facts concerning what there is unrestrictedly are contingent or necessary, and this dispute is orthogonal to their dispute over what in fact exists unrestrictedly.

[48] I do not mean to suggest that the non-Lewisian cannot consistently claim that such facts are necessary, merely that it would be very strange for them to do so. Even those who think that every (unrestricted quantifier) thing that exists necessarily exists (for example Timothy Williamson [149]) hold that for every (unrestricted quantifier) concrete thing it is contingent that that thing both exists and is concrete, and hence agree that facts concerning what concrete beings there are (unrestrictedly) are contingent.

[49] If it is possible that p then it is possibly necessary, hence, given B, true.

doesn't require causal contact with the subject matter. It is a principled boundary, though motivated within the very modal realism that is in dispute. (I am mounting a defensive operation, and will be content with a standoff.)"[50]

The question I want to ask is this: is Lewis right to be mounting solely a defensive operation? Should he be content with a standoff? And in general, is it methodologically sound that Lewis has responded to both the modal irrelevance objection and the epistemological objection by relying on tenets of his theory of modality? My answer to each of these questions is no, and below I will explain why.

5.3.3 Epistemology and the Integration Challenge

Let me return to the epistemological objection. The objection is basically an instance of Peacocke's Integration Challenge[51], itself a generalisation of Benacerraf's dilemma for the philosophy of maths.[52]

Peacocke's guiding thought is that it is no good to give a story of the metaphysics of a discourse which makes it an utter mystery how we obtain knowledge of the truths in that discourse: our metaphysics must be integrated with our epistemology. This demand can come in a weak and a strong version. The weak version demands only that there be *some* epistemological story that can be told about how we come to know (some of) the truths of a certain discourse, given the metaphysics of that discourse that have been proposed. The strong version, which is the version Peacocke advances and which I will accept here, demands that we be able to make sense of (what appears to be) our *actual* methods of coming to know the truths of that discourse, given the proposed metaphysics. The stronger demand resists a revisionary epistemology. It is no good to be given a metaphysics of a discourse D and told that we are able to obtain knowledge concerning the truths of that discourse provided we change our practises. We believe ourselves to have had knowledge of some of the truths of D prior to being offered this metaphysical picture of the D truths, and what we want is an epistemological story that allows us to have had such knowledge. The metaphysics must integrate with what has in fact been our way of finding out about these

[50] [ibid. p112]
[51] [102]
[52] [14]

truths.

In my opinion, any proposed analysis of modality must meet the strong integration challenge. But crucially, I claim that it must do so *independently* of the analysis in question. That is to say, it is not sufficient for an analysis to meet the challenge by showing that *according to the analysis* there is no mystery that our actual method of finding out about Φ-facts is successful; we must be given reason independent of the analysis in question to think that our actual method of finding out about Φ-facts would be successful if the metaphysics of the Φ-facts is as the analysis says it is.

Why this constraint? Because the fulfilling of the integration challenge is a necessary requirement for the acceptability of any analysis. Warrant for an analysis of Φ in terms of Ψ requires warrant for the claim that knowledge of Ψ-facts is obtainable by the method we in fact use to gain knowledge of the Φ-facts. So someone who is not a Lewisian realist can only obtain warrant for believing Lewisian realism if she obtains warrant for believing that our method of obtaining modal knowledge could lead to knowledge of non-actual individuals. But it is no good if the latter claim itself relies on Lewisian realism, since it is the truth of that very theory that is up for debate. If someone is unconvinced about the truth of Lewisian realism on the grounds that it makes a mystery of modal knowledge, she is obviously not going to be assuaged by appeal to a claim that only the Lewisian realist believes; such a claim is part and parcel of the very theory she is doubting.

If it is only according to Lewisian realism that knowledge of non-actual individuals does not require causal interaction with those individuals, then warrant for thinking that the integration challenge is met presupposes warrant for Lewisian realism. In that case the truth of Lewisian realism is being presupposed in a way that is viciously circular. The non-Lewisian is being presented with a theory the acceptability of which relies on a claim that presupposes the truth of that very theory. In that case they should not be moved to accept either claim. If I have no belief in either of two claims p and q, and if warrant for each presupposes warrant for the other then, I claim, I should not be moved to believe either: warrant for either claim can never get off the ground.

Lewis responds to the epistemological objection by mounting a defensive operation: by showing

that by the lights of the Lewisian realist there is no problem concerning knowledge of what goes on at non-actual Lewisian worlds. But Lewis *shouldn't* simply be mounting a defensive operation. The challenge is not, properly understood, that modal knowledge is a mystery *by his own lights*; the challenge is that we, those who have not accepted Lewisian realism, have no reason to think that modal knowledge is anything but a mystery if Lewisian realism gives a correct analysis of the modal. As such, since an analysis of modality is only acceptable if we have reason to think that it gives a satisfactory story about modal epistemology, we have no reason to accept Lewisian realism. The non-Lewisian realist has no reason to think that Lewisian realism is compatible with an acceptable modal epistemology, and thus has no reason to accept it as an analysis of the modal. So it seems that we have no reason to think that Lewisian realism can be held side by side with an adequate account of modal knowledge; hence, it should not be accepted.

What Lewis needs to do to respond to the epistemological objection is provide reason *independent* of the truth of Lewisian realism to think that one can have knowledge about non-actual possible worlds; reason that does not presuppose doctrines that the non-Lewisian is going to reject, such as the claim that what concrete objects unrestrictedly exist do so as a matter of necessity. That is what Lewis cannot give, seemingly; and so while he may have argued that there is no epistemological problem by his own lights, he has given us no reason to think that there is no epistemological problem by *our* lights. Since we would need to be convinced that there is no epistemological problem in order to find Lewisian realism an acceptable analysis of the modal, this is a good reason not to accept Lewisian realism.

5.3.4 Inference to the best explanation

Let me address a possible worry. The worry is that in insisting on having a reason *independent* of Lewisian realism to accept that one could have knowledge about non-actual worlds I am somehow denying inference to the best explanation. Ontological hypotheses should be accepted if they have eminent utility[53], so the objection goes, and so we should accept the Lewisian realists' account of how we can have knowledge about non-actual worlds if Lewisian realism has eminent

[53]C.f. Lewis [66, pvii] "systematic philosophy goes more easily if we may presuppose modal realism in our analyses. I take this to be a good reason to think that modal realism is true, just as the utility of set theory in mathematics is a good reason to believe that there are sets."

utility. Whether or not there is a reason independent of Lewisian realism to accept the account is neither here nor there; we have reason to accept Lewisian realism if the theory is sufficiently beneficial, and I have done nothing to show that it isn't.

I do not wish to deny the legitimacy of inference to the best explanation; I wish only to insist on a particular understanding of how we judge a theory to be the best explanation of some phenomena. There is a view of inference to the best explanation that says that we should accept a theory T if, *according to the lights of T*, T provides the best explanation of some relevant phenomena P. Well if that is what inference to the best explanation is then I do indeed deny it. Why on Earth should I care about whether or not a theory provides a good explanation by its own lights when I do not believe that the theory is true? What I should care about is whether or not a theory provides a good explanation by the lights of the theory I currently hold, or at least a theory that is as neutral as possible between the one I currently hold and its proposed rival. But Lewisian realism only provides a good explanation of the modal according to its own lights since it is only according to the lights of Lewisian realism that one can have knowledge of concreta without being in some causal relationship with those concreta; hence, by the lights of anyone other than the Lewisian realist, Lewisian realism makes modal knowledge an utter mystery, and thus is not a good explanation of the modal. By its *own* lights Lewisian realism doesn't make modal knowledge a mystery since Lewisian realism implies the falsehood of certain of the claims which are relied upon to show that we cannot have a priori knowledge of concreta (namely, the contingency of what concreta unrestrictedly exist). But so what? Unless I have reason to accept Lewisian realism I have every reason to make those claims, and hence every reason to think that Lewisian realism does not give an adequate account of modal knowledge.

I cannot bring myself to accept a theory if that theory has sufficiently bad consequences by the lights of my current theory. It is not enough that if the new theory is true it does not have those bad consequences; why should I care about that when I currently have no reason to think that the theory *is* true? So much the worse for the claim that we judge theories on the basis of the explanation they provide on the assumption that they are true; rather we should be concerned with the explanation they provide as judged by the lights of the theory we currently hold. I am not going to accept that Lewisian realism provides any sort of explanation of the modal unless I

have reason to think it does not render modal knowledge a mystery; now granted I have reason to think that *if* Lewisian realism is true then there in fact is no such problem — for I can see that *according to Lewisian realism* there is no problem with how we can come to have knowledge of non-actual concreta — but since the antecedent of that conditional is precisely what is up for debate, this is not going to move me to accept the consequent.

What I deny, then, is that when deciding between rival theories we suppose each of the rivals to be true, weigh up the benefits and the costs of accepting the theory on that assumption, and then adopt the theory that maximises benefit and minimises cost. Why should I be interested in how a theory fares as to costs and benefits on the assumption that that theory is true? Should I not be interested in how it fares by my current lights? After all, the theory I currently hold is the theory I believe, and believe I have good reason to believe, is true. Why should I care how a theory fares on an assumption that I think is false? Shouldn't I ask how it fares on an assumption I think is true? When deciding between rival theories I should weigh up their costs and benefits as they are by the lights of my current theory, and by this method Lewisian realism is not very good at all, because it fails to yield an adequate account of modal knowledge. I conjecture that the reason philosophers have sometimes been attracted to the methodology I am denying is that they think that otherwise abandoning one theory in favour of another would never happen. For why, if I am judging rival theories by the lights of my current theory, would I ever abandon my current theory, since obviously any rival is going to be false by the lights of my current theory. (That's what it is to be a rival.) But really this is no problem. It is true that any rival theory is going to be false by the lights of my current theory, but that does not rule it out being better *justified.* A rival theory, although false by my current lights, will be better justified than my current theory if it provides a better explanation for data which is recognised by my theory as requiring explanation, without violating what are, by my current lights, constraints on what can count as an explanation for that data. It is clear then how theory progression is possible; we are justified in abandoning our current theory T and adopting T* when T* is, by the lights of T, more explanatory than T, even though T* is false by the lights of T. So I can be justified in accepting a theory which is to my lights false, because it is, by my lights, more explanatory. The idea that theories are chosen by weighing up the rivals by their own lights is, I think, completely

hopeless. For since the means by which we measure the costs and benefits of rival theories are themselves determined by our theory, this would make it a simple matter for someone to force on us an unacceptable theory that we could not then get rid of. For methodology is but a part of our total theory, and if a rival theory were presented to us which said of itself that it is the best theory and cannot be improved upon then, if we were to judge that theory by its own lights, we would indeed be convinced that it was the best theory and that it could not be improved upon.[54] That proves, I think, that we must, when judging rival theories, keep fixed some claims our current theory makes. In particular we must keep fixed claims concerning how theories are to be judged. Similarly, we must keep fixed claims concerning what the data is that needs explaining: an atheist is obviously not going to be impressed by the argument that they should believe in God because the existence of God is the best explanation for the existence of heaven — the existence of heaven is not a datum that the theist recognises as needing an explanation. If the theist wants to convert the atheist they must rely only on data that the atheist will recognise by their own lights.

That much is I think relatively uncontroversial: that certain claims need to be kept fixed when deciding whether or not to revise one's theory; that we shouldn't allow ourselves to accept a rival theory on the grounds that according to it these claims are false. All I am urging is an extension of this very natural line of thought: that we must also keep fixed certain claims concerning how data can be explained. In particular, when judging theories with rival ontologies, it seems appropriate to me to keep fixed the constraint that knowledge of concreta can only be explained if one has causal interaction with those concreta. This is why Lewisian realism does not provide an adequate explanation of how we can have modal knowledge.

5.3.5 Irrelevance again

I have argued that Lewis has not provided an adequate response to the epistemological objection because he presupposes warrant for Lewisian realism in answering a challenge that has to be met in order for Lewisian realism to be warranted in the first place. In this section I will make

[54] As an example, consider a parody of the Holy Gospels which proclaim themselves to be the truth, the whole truth, and nothing but the truth, as divinely revealed by God.

a similar case regarding his response to the modal irrelevance objection. The idea is this.[55] The acceptability of an analysis requires at least that the analysis be materially adequate. In that case, warrant for an analysis of Φ in terms of Ψ presupposes warrant for the claim that for every Φ truth there is a corresponding Ψ truth. If warrant for this latter claim is lacking we have no reason to think that the truths about the Ψs provide an adequate reductive basis for truths about the Φs — the truths about the Ψs will appear *irrelevant* as an account of what is true of the Φs. In the case of Lewisian realism, then, warrant for the claim that possibility is to be analysed as what is true at some world presupposes warrant for the claim that for every modal truth there is a corresponding truth concerning what is true at some world. And because the challenge is to show how one can attain warrant for Lewisian realism, justification for the required claim must be given which does not presuppose the truth of the theory in question. So we must be given reason to think that there are enough truths concerning what occurs at some Lewisian world to provide an adequate reductive basis for the modal which is independent of the truth of Lewisian realism. But that is what Lewis cannot give us. The modal operators might be quantifiers over possible worlds *by his lights*; but that is only true according to the very theory that is up for debate. If we are to think that facts concerning Lewisian worlds are relevant to modal facts — to think that such facts can serve as a reductive basis for the modal — we must be shown that by *our* lights, or at least by the lights of a neutral theory, there is the required correspondence between what is true at some world and what is possible. But we will never be so convinced, for any reason to think that there is such a correspondence presupposes that there are the plurality of worlds Lewis believes in, and the only reason to believe in such worlds is if you think that there is a Lewisian world for every possible way the actual world could have been — which is only true according to Lewisian realism!

Let me make the objection more explicit. In order to make us abandon our worry that facts about Lewisian worlds are not the right type of facts to allow for a reduction of the modal, Lewis has to convince us that there is a Lewisian world for every way the actual world might be. That there is such a totality of Lewisian worlds is implied by Lewis' principle of recombination (Com), stated again below:

[55] I refrain from commenting on how close I think this is to the original intent behind the objection as put forward by van Inwagen and Jubien.

Com: For any wholly distinct things $x_1, x_2, ...x_n$ there is a world containing any positive number of duplicates of each, and no thing which does not overlap any of those duplicates, size and space permitting.

And the problem is this: one would only be inclined to believe Com if one was already a Lewisian realist. Suppose we were not interested in modality but doing ontology for its own sake. Com would not look remotely plausible as a principle about what maximal mereological sums of spatio-temporally related individuals existed. The only warrant one can have, or so it seems to me, to believe that Com tells us about such mereological sums is that Com tells us about possibilities, and that such mereological sums simply *are* maximal possibilities. But in that case Lewis is using the truth of Lewisian realism to justify the truth of that very theory. For one will only accept Lewisian realism if one believes that there is a Lewisian world for every way the actual world might be; otherwise we cannot overcome the worry that facts about Lewisian worlds are the wrong type of facts to reduce modal facts to. But one will only believe that there are enough Lewisian worlds for this if one believes Com, and one will only believe Com if one believes that every way the actual world could be is a way that some Lewisian world is; i.e. only if one is *already* a Lewisian realist. There is a methodological circularity then: Lewis is using his theory to defend his theory; and hence we have no reason to believe it.

The problem is nothing to do with the principle of recombination in particular, of course. The problem would arise with *any* principle that tells us that there are enough Lewisian worlds for facts about them to serve as an adequate reductive basis for modal facts. By definition, we can only have causal interaction with (at most) one Lewisian world; hence (assuming that we can only have empirical justification that something exists if we can stand in some causal relation with that thing) we cannot have empirical justification of the existence of more than one Lewisian world. But we need to be justified in thinking that there is more than one Lewisian world in order to have a justified belief that facts about Lewisian worlds serve as an adequate reductive basis for facts about modality, since we know that there are non-actual possibilities. So it is a necessary pre-condition for a justified belief that facts about Lewisian worlds serve as an adequate reductive basis for modal facts that we have an a priori justification that there is more than one Lewisian world. But a priori justification that there is more than one Lewisian

world only seems possible if we have prior a priori justification that there is a Lewisian world for every way the actual world might be. Then we can infer that there is more than one Lewisian world from our a priori knowledge that there are non-actual possibilities. In that case warrant for Lewisian realism necessarily presupposes warrant for itself. We are only justified in believing Lewisian realism if we are justified in thinking that there is more than one Lewisian world. Justification that there is more than one Lewisian world must be a priori. But we can only have a priori justification that there is more than one Lewisian world if we have a priori justification that there is a Lewisian world for every way the actual world might have been — i.e. only if we have a priori justification in Lewisian realism.

So the objection, in sum, is this. Lewis asks us to believe in the existence of all these cosmoi because of the utility they would have if they exist. Of particular interest is that they would allow for an analysis of the modal in terms of the non-modal. But, so the objection goes, we only have reason to think that facts about Lewisian worlds serve as an adequate basis for a reduction of modal facts if we already believe that there are Lewis-many such worlds. So I should only accept the reasons to believe in the plurality of Lewisian worlds if I already believe in them.

And again, let me make clear that I am by no means denying the legitimacy of inference to the best explanation here. My claim is simply that Lewisian realism can only be said to be an explanation of the modal if there is the requisite correspondence between facts about worlds and modal facts; since I have no reason to believe such a correspondence holds I have no reason to think that Lewisian realism offers an explanation, let alone the best explanation, of the modal. While I accept that Lewisian realism might provide the best explanation of the modal from the point of view of the Lewisian realist, that is precisely because they deny the claims that are true by the lights of my current theory, or indeed by any neutral theory, that rule out Lewisian realism as being the best explanation of the modal.

5.3.6 Methodology: Lewis versus Forrest

I have argued that one could only be warranted in accepting Lewisian realism if one has reason *independent from Lewisian realism* to think that there is a Lewisian world for every way the actual world might be, but that such reason will not be forthcoming. In order to illustrate the

objection against Lewis, and to show why it does not generalise to knock out any informative reduction, let us look at a rival theory of modality which I think avoids this particular objection admirably. That is the theory offered by Peter Forrest in his [42]. Forrest's theory endorses the scheme

$$WN : \Diamond p \leftrightarrow \text{p is true under}^{56} \text{ some world nature.}^{57}$$

The world nature of a world W_1, for Forrest, is the property which is the conjunction of all the natural properties of W_1. World natures, then, are Armstrongian structural universals — albeit uninstantiated as well as instantiated ones — which are "in some way *composed* of simpler properties and relations."[58] Forrest recognises that if WN is to prove to be an acceptable reduction of modality then he has to convince us that there are enough of these structural universals to ensure that every Lewisian world has (or rather, would have, were they to exist) a world nature, otherwise we would not be able to account for the possibility of some propositions which are intuitively possible.

So in order for Forrest's proposed reduction to be acceptable we need to be convinced that what is true under some world nature does not come apart too much from what we pre-theoretically think is possible. Forrest provides us with various principles to tell us about what structural universals there are. Crucial among these is the following principle, which I will call Gen.

> Gen: "If R is an m-adic property or relation and S is an n-adic one, then R×S is just the (m+n)-adic relation which holds between $x_1, ..., x_m, y_1, ..., y_n$ just in case R holds between $x_1, ..., x_m$ and S holds between $y_1, ..., y_n$."[59]

Gen is Forrest's substitute for Lewis' Com. It generates structural universals by an operation on simpler properties and relations, just as Com generates more complex individuals by an operation on simpler individuals (i.e. maximal mereological sums from non-maximal mereological sums). Let us grant for the sake of argument that Gen (in conjunction with certain other principles that Forrest appeals to) implies the existence of a plurality of structural universals which

[56] 'true under' is Forrest's replacement of Lewis' 'true at' relation.
[57] ibid. p17
[58] [ibid. p17]
[59] [ibid. p17]

makes what is true under some world nature cohere sufficiently with what we pre-theoretically think is possible. As before, the question is whether or not Gen is true. And the benefit of Forrest's theory over Lewisian realism is that Gen looks like exactly the sort of principle we would accept if we were interested in a theory of structural universals for its own sake rather than as a proposed reduction of modality. Gen exploits the analogy between complex properties constructed from simpler properties and complex mereological sums constructed from simpler individuals that believers in structural universals tend to accept. Many people believe that if we are going in for talk about mereological sums then we ought to believe that given any individuals a and b there is an additional object which is the mereological sum of a and b. The parallel thought is that if we are going in for talk about structural universals then we ought to believe that given a property G and a property H then there is a property G×H which holds between x and y just in case x has G and y has H.[60] It is this thought that Forrest is aiming to capture with Gen.

Gen does not rely for its plausibility on the supposition that the complex properties which are the structural universals are maximal possibilities; thus Forrest does not need to appeal to the truth of WN in order to justify Gen. And hence Gen can be used to justify WN without falling into the methodological circularity that befell Lewisian realism. That means that we can recognise that Forrest's reduction of modality is a *realist* reduction. We don't need to believe WN to be convinced that the reduction it offers is successful. If we believe in structural universals then we can be convinced *independently of the theory that implies WN* that there is necessarily a one-one correspondence between what is possible and what is true under some world nature. Thus we can be convinced, from the standpoint of the theory that we currently accept, that facts about what is true under some world nature serve as an appropriate reductive basis for modal facts, and hence abandon any worry that Forrest is in fact an eliminativist about modality. So no modal irrelevance objection applies to Forrest.

I have argued that we should only accept a theory T which reduces modal facts to some realm of facts F if we have grounds *independent of theory T* to think that F serves as an adequate reductive basis for the modal. This is what Forrest offers us when he offers us reason to believe

[60] cf ibid. p17

that what is possible is true under some world nature which are *independent* of the success of his reduction of the modal; it is what is lacking in Lewis.

Forrest sees a similar point: when anticipating the objection that his theory makes use of a controversial ontology (not only structural universals but uninstantiated ones) he says "All this is, no doubt, controversial. But in many cases I provided an *independent* motivation for what I require."[61] This is the crucial difference between Forrest's theory and Lewis'. If you believe in mereological sums then you will not accept Com as a principle which tells you what maximal mereological sums of worldmates exist unless you already believe that such things are maximal possibilities. Whereas if you believe in complex properties then I think you are likely to accept Gen as a principle that tells you about what structural universals exist before you even consider whether such things are useful to modality. Note that I am not here arguing for Forrest's theory; personally I do not believe it because I do not believe in uninstantiated universals (and I am agnostic on instantiated ones). I merely aim to praise Forrest's methodology — whether his theory is successful is another matter.

5.3.7 Conclusion

Let us return to the charge Chihara leveled against Lewis concerning the modal irrelevance objection. The charge was that Lewis can only respond to the modal irrelevance objection by assuming the truth of Lewisian realism: the very theory that is up for question. I said that the question as to who has won the debate hinged on whether or not Lewis is entitled to respond to the objection from his own perspective — the perspective of the Lewisian realist. Well, in a sense he is and in a sense he isn't. He is entitled to so respond in so far as he is attempting to justify the claim that *by his own lights* the modal irrelevance objection does not show that his reduction is unsuccessful. On this score we should agree with Lewis; the modal irrelevance objection should not be expected to convince the Lewisian realist that his reduction of modality does not work. But in so far as Lewis is trying to give us *reason* to be a Lewisian realist he is not entitled to respond to the modal irrelevance objection from the point of view of the modal realist, since that response is obviously one his opponent will not accept.

[61] [ibid. p24]

We can see also how Jubien's two objections are related. It is the lack of reason to believe that there is a Lewisian world for every way the actual world might have been that gives rise to the modal irrelevance objection, because it is the absence of there being enough facts about how some Lewisian world is that makes facts about Lewisian worlds the wrong type of facts to serve as an adequate reductive basis for the modal.

The modal irrelevance objection, as I have developed it, should not be understood as an attempt to expose some internal inadequacy of Lewisian realism; rather it should be seen as a methodological objection to Lewisian realism. It challenges the justification for accepting Lewisian realism by showing that one should only accept that the theory is adequate if one already believes the theory. One is justified in thinking that Lewisian realism provides an adequate reduction of the modal only if one is justified in thinking that facts about what is true at some Lewisian world are the right kind of facts to serve as a reductive basis for modal facts. One is justified in thinking *that* only if one is justified in thinking that there are Lewis-many Lewisian worlds. But the justification for the claim that there are Lewis-many Lewisian worlds relies on the principle of recombination, which one is only justified in believing if one is a Lewisian realist. In that case one can only have justification for accepting Lewisian realism if one is already a Lewisian realist; which makes justification for Lewisian realism unobtainable.

The modal irrelevance objection, as I have construed it, and the epistemological objection, are different; but there are important similarities between them. The problem the epistemological objection raises for Lewis is that we have no means of knowing what goes on at spacetimes other than our own, even if there are any, thus making knowledge of non-actual possibilities impossible on the Lewisian account. Lewis attempts to answer us by showing that if he is right and the unrestricted domain of concreta cannot vary then there is no need for a causal requirement on knowledge of what concreta unrestrictedly exist. But that will not help someone who thinks that Lewisian realism is unacceptable on the grounds that it makes modal knowledge a mystery, for it presupposes the truth of the theory that is up for debate. What we want from the Lewisian is an *independent* assurance that we can know about what goes on at these worlds. The non-Lewisian-realist will only find Lewisian realism acceptable if they have such an assurance, but it seems that the Lewisian cannot give it because there is no independent route to knowledge of what

goes on at Lewisian worlds: no route that does not proceed via the assumption that something is true at a Lewisian world iff it is possible. And it is precisely this lack of independent access to Lewisian worlds that forms the backbone of my modal irrelevance objection. For the analysis of possibility as truth at a world to be acceptable, I argued, we must have warrant for thinking it materially adequate that does not presuppose the truth of the theory in question. That is what I argued we cannot have; the lack of independent access to non-actual Lewisian worlds means that the only grounds for beliefs concerning them would be based on our modal beliefs and a belief in the correspondence between modal truth and truth at a world. Thus there is no independent means to justify the Lewisian analysis; no way to provide warrant for that theory that does not presuppose the truth of that theory.

So the lack of independent access leads to two problems for Lewisian realism: firstly, that we cannot know what goes on at non-actual Lewisian worlds, and secondly, we cannot even know that there are any. The first problem results in the epistemological objection to Lewisian realism: that the Lewisian has given us no independent reason to think that on their account modal knowledge is anything other than utterly mysterious. The second problem results in the modal irrelevance objection: that the Lewisian realist has given us no independent reason to think that there are any non-actual worlds, and hence no independent reason to think facts about Lewisian worlds serve as an adequate reductive basis for the modal. The Lewisian realist can, of course, give us theory laden reasons to resist both the problems. On the assumption that Lewisian realism is true it follows that we can have knowledge that there are non-actual Lewisian worlds, and knowledge about what goes on at them. But that is beside the point when the epistemological and irrelevance objections are understood as they should be: as external objections, challenging our warrant for accepting Lewisian realism. What is needed to answer them is some reason to think that we can know about these worlds that *does not* presuppose the truth of Lewisian realism. That is what we cannot have, and that is why Lewisian realism is unacceptable.

Lewisian realism provides an adequate reduction of modality on the assumption that it is true. But why should we care about that? Why should I care about the benefits of a theory on the assumption that it is true when I do not myself believe that it is true. And in particular, how

can I be expected to accept this as a reason to believe Lewisian realism? I am being asked to accept that Lewisian realism is true on the grounds that it is explanatory if it is true. But I have no grounds to believe that it is explanatory: for to have grounds I would have to discharge the antecedent of the previous conditional. But the antecedent is precisely what is up for debate: whether or not Lewisian realism is true.

5.4 Modal Agnosticism

Orthodox Lewisian realism has to be abandoned on epistemological grounds, but there is a theory available which accepts the Lewisian analysis of modality, warrant for which requires neither warrant for believing in non-actual worlds nor warrant for believing propositions concerning what goes on at non-actual worlds (subject to an upcoming qualification). This is the modal agnosticism recently put forward by John Divers.[62] Divers' modal agnostic asserts the Lewisian analysis that P is possible iff it is true at some world, but refuses to assert anything that entails that there is more than one world. But she also refuses to assert that there is just one world. Such an agnostic, then, refuses to make claims asserting the mere possibility of propositions (with the possible exception of some de re possibility claims, see below); but they *are* prepared to make necessity claims. The idea is that while the agnostic refuses to assert anything *positive* about what goes on at other worlds, they are prepared to make negative judgements; i.e. to rule out certain things. So the agnostic will refuse to assert that there is a world with a talking donkey, but they're perfectly prepared to deny that there are any worlds that both do and do not contain talking donkeys. The position is meant to be analogous to van Frassen's position regarding unobservables: as a consequence of his belief that we have no warrant to think that there are any unobservables van Frassen will refuse to assert anything he thinks entail their existence; but while he will not assert (nor deny) that there any positively charged unobservables, he is perfectly prepared to deny that there any unobservables that are both positively charged and not positively charged. So just as van Frassen will assert negative existentials concerning unobservables but not positive existentials, so the modal agnostic will make negative existentials concerning what non-actual worlds there are but not positive existentials.

[62][33]

The agnostic, then, can make necessity claims, since they on the Lewisian account are negative existentials and do not entail the existence of any non-actual worlds (necessarily p iff at all worlds p, i.e. if at *no* world ¬p), but not claims of mere possibility, since they on the Lewisian account are positive existentials and do have such entailments (possibly p iff *there is* a world at which p). They can also make counterfactual claims, since they are negative existentials (had it been that p it would have been that q iff *there is no* selected world at which p and ¬q), but not might counterfactuals since they are positive existentials (had it been that p it might have been that q iff *there is* a selected world at which p and q). The agnostic, if they are a counterpart theorist who accepts intra-world counterparts (see section [3.1.4]) can even make some claims of mere possibility; for if I have an actual counterpart who is Φ then that allows me to assert that I could have been Φ without committing myself to the existence of non-actual worlds. What the agnostic will not be able to assert are claims of de dicto mere possibility, or de re possibilities of a thing x such that x has no actual counterparts which are that way.

This seems to avoid the above problems. The agnostic does not commit themselves to the claim that there are non-actual worlds; a fortiori they do not face the demand to justify this claim independently of their analysis, and hence they face no modal irrelevance objection. Likewise, while the agnostic lays claim to have knowledge about what goes on at non-actual worlds, it is all negative knowledge: knowledge about what *doesn't* happen at any Lewisian world. They make no positive claims about what *does* happen at any non-actual world. This seems to avoid the epistemological problem, as it is positive knowledge concerning non-actual Lewisian worlds that seems mysterious, and which we would rightly demand independent reason to think we have. Negative knowledge, that no Lewisian world is such and such a way, seems far more tractable. Divers gives the following metaphor to make this point.[63] Suppose you are, and have always been, stuck on a desert island, and all you can see around you is sea. Your island has a single palm tree on it, and you wonder whether there are any islands so rich in resources as to have *two* palm trees. You cannot know whether or not there is such an island, seemingly. Your isolation from land other than your own small island means that you cannot even know whether or not there *are* other islands, let alone whether or not there are any with two palm trees; such knowledge would require interaction with islands beyond your own which you sadly lack. But

[63] ibid.

what you *can* come to know, seemingly, is that there are no islands which both have two palm trees and have no palm trees. While causal interaction with other islands is a necessary condition for the acquisition of knowledge that there is an island such that p (when your own is such that ¬p), a priori logical reasoning is all that is needed for knowledge that there are no islands such that both p and ¬p. So while we should be dubious about any claim you make to have positive knowledge about what some other islands are like, since it would be an utter mystery how you came to have such knowledge, we should be happy with your claims to knowledge about what *no* island is like. The knowledge that no island is such that p and ¬p is not mysterious, for it follows from the law of non-contradiction which is a priori and which therefore we do not require empirical evidence for. The agnostic is in the same position as the marooned logician; it is no mystery how they can obtain the negative knowledge that no world is such that p and ¬p — at least it is no more a mystery than our knowledge that at the farthest reaches of the universe there is no planet which is both inhabited and not inhabited; what would be mysterious is knowledge that some world *is* such that p (when our own is not), but the agnostic makes no such claims.

So the agnostic faces neither of the two objections I take to be sufficient to reject Lewisian realism. Are there any other reasons to reject agnosticism? The reason I cannot accept Lewisian agnosticism is simply due to a deep conviction that there are facts about how the world could have been but isn't. I do not even need to commit myself to the claim that they are knowable in order to pose the problem for the agnostic. The agnostic leaves it as an open question that every truth is necessary; but I do not think this question is open. I think I know that some falsehoods are possible, even if I do not know *of* any falsehood that it is possible. Even if I could live with the fact that for every proposition it is an open question whether or not it is necessary, I cannot live with the fact that it is an open question that every proposition is necessary. But the de dicto knowledge that the world could have been otherwise is all that is needed to reject agnosticism.

In rejecting agnosticism for this reason I am refusing to play the game Divers wants us to play. Divers wants the opponent of the agnostic to show us what *practical* difference there is between that agnostic and the non-agnostic. That is, he wants the opponent to point to some case in which those who are prepared to assert what the agnostic is not will act differently from the agnostic. The idea being, I take it, that unless we can point to the *function* of making

non-actual possibility claims we can have no principled objection to a position which refrains from making such claims. But I take investigation of modal reality to be worthwhile in its own right, not merely subservient to some other purpose. If one thought that the sole reason we should engage in modalising was to serve some function F then, provided an account of modality let us serve this function, it would be acceptable. But it seems to me that the primary function of modalising is simply to discover how the world might have been. If it is true, as I believe it is, that the world might have been some other way than it in fact is, then it is indeed the case that the agnostic cannot carry out every function of modalising the non-agnostic can.

Furthermore, I simply cannot see how the agnostic could ever acquire warrant for the Lewisian analysis in the first place. Given that we believe that the world could have been otherwise then, given the agnostic's refusal to commit to the existence of non-actual worlds, why would the agnostic ever think the Lewisian analysis of modality a correct analysis? An analysis of Φ will only ever look attractive if, by and large, what comes out as being Φ according to the analysis coheres with our beliefs as to what is Φ independently of the analysis. So how could one ever accept the Lewisian analysis if they refused to believe that there were non-actual worlds, given the universal conviction that there are ways the world could have been that it is not? I can understand how someone could come to give up the conviction that the world could have been otherwise were they in a position whereby they had already accepted the Lewisian analysis, but I can't understand why they would have accepted the analysis in the first place.

Chapter 6

Conclusion — Quasi-Conventionalism

Where do we stand? I have attempted to locate the source of modal truth in the actual by looking for actual entities that might be truthmakers for modal truths, and I have not found anything. I have considered the view that we should ground the modal in the existence of non-actual ontology, but I have rejected the existence of any such things. What remains for one who does not want to accept brute modal facts, but who nevertheless wants to hold that we (at least sometimes) speak truly when we modalise?

Consider again the deflationist story I accepted in chapter 2. That was, in Putnam's term, a *quasi-conventionalist* story about the necessary a posteriori and about the essential/accidental properties distinction. There we did not appeal to any brute facts concerning the natures or essences of entities or kinds to explain these a posteriori necessities or essentialist claims, but we did allow an underlying realm of modal facts — facts concerning the qualitative ways the world could have been — which, in addition to the way we use names and natural kind terms, accounted for the truth of those claims.

The deflationist view at that point is that there is a space of mind-independent modal facts that represent, contra Kripke, a qualitative way for things to be. But those facts leave it open

what sentences come out as necessary or contingent, for they do not settle whether or not in a particular world the watery stuff represents water, or the gold-like stuff gold, or a particular human Aristotle. *We* settle such questions, as a matter of the conventions governing usage of the terms 'water', 'gold' and 'Aristotle'. It is these underlying modal facts, these facts concerning the qualitative way the world could be, whose truth I am trying to ground, and which are proving to be elusive.

But maybe we should go even further, and apply the deflationist's trick again to do away with the commitment to an underlying realm of modal facts. The deflationist strategy is to grant that there is some acceptable underlying realm of mind-independent facts and show how these facts together with conventions governing the use of language in some target discourse result in assertions in the target discourse coming out true. Deflationism about modality, then, would grant that there is some underlying realm of *non-modal* mind-independent facts and show how these facts together with the conventions governing modal language result in modal statements expressing truths.

This is the position that I want to explore in this final chapter. I will only be able to give a sketch of the theory, but I hope to be able at least to give an idea for how someone might agree with everything I have said so far but resist the pressure to give up on modal talk. It will be helpful to approach the position in question by first considering typical objections to two theories.

Ersatzism.

Consider (one of) Lewis' main objection(s) to ersatzism about possible worlds (the view that possible worlds are actually existing abstract representations of ways for the world to be): the objection is that the ersatzist must take some modal notion as primitive.[1] Why is this? Well in general, whatever the ersatzist thinks worlds are she will think that there are entities corresponding to impossible worlds as well as entities corresponding to possible worlds. If worlds are sets of propositions then there are sets of propositions that could not all be true together just as there are sets of propositions that could all be true together; if worlds primitively represent some counterfactual situation as obtaining then there will be ones that represent impossible situations

[1] See especially [66, Ch.3].

as obtaining just as there are ones that represent possible situations as obtaining. In that case the ersatzist cannot claim that p is possible iff it is true at some world, they can claim only that p is possible iff it is true at some *possible* world. And, so the objection goes, the ersatzist has nothing to say about what makes some worlds possible and others not; it simply *just is the case* that the situations said to obtain by some worlds could obtain and that the situations said to obtain by others could not. Contrast this with the Lewisian who thinks that being possible is just being true at a world (not a possible world), because there simply *are no* worlds according to which impossibilities hold true. It is a powerful objection, because there is something deeply suspicious about primitive modal facts; the same thing that is suspicious about primitive tensed facts. Just as the present only seems to ground the truth of how things are now, and not how they were or will be, so does actuality seem to ground only the truth of how things in fact are, not how they could have been.[2]

Conventionalism.

The conventionalist wants to locate the source of modal truth in language, not in the world. Traditionally, the conventionalist claims that the necessary truths are those which are true by convention (or true in virtue of convention). But this theory is hopeless. What is a matter of convention is that sentences express the propositions they do, for we could have used the same sentences to say something different. So if a sentence S expresses a truth then in this sense it does so in virtue of convention; we could have used S to express a falsehood. But in this sense *every* true sentence is true by convention; no distinction is carved between the necessary and the contingent. If such a distinction is to be drawn then the sense of 'true by convention' the conventionalist is interested in must not simply be the trivial sense in which all sentences are true by convention; rather, the conventionalist must hold that there is a notion of 'true by convention' such that not only is it a result of convention that a sentence S expresses a truth but that it is a result of convention that the truth expressed by S is true. So while a contingent sentence S and a necessary sentence S* are both 'true by convention' in the sense that we could have used those sentences to express falsehoods, the truth expressed by S* but not the truth expressed by S holds in virtue of convention. It is in this sense that the necessary truths are true by convention.

[2]This is Sider's criticism of brute modal truth, which I accepted at the outset in this thesis.

But that is simply false. It is not true by convention in this sense that 2+2=4 or that all bachelors are unmarried etc; those claims are true in virtue of what numbers and bachelors are like. Every sentence is such that the fact that it expresses the proposition it does is a matter of convention, but every true proposition is true in virtue of how the world is. The *only* propositions that we could legitimately claim are true by convention in this required sense, then, are those that say something *about conventions*: e.g. that there are conventions, or that such-and-such a convention holds, etc. *Those* propositions are true by convention, in the same way that the truth that there are tables is true in virtue of the tables. So it is not the necessary propositions that are true by convention in this sense; it is this small subset of the contingent propositions that are *about* conventions.

So traditional conventionalism must be abandoned, and traditional ersatzism does not give an adequate answer to the source of modal truth. But the ersatzist need not admit of primitive modality, and the conventionalist need not commit themselves to truth by convention in the objectionable sense. Perhaps there are no modal facts and the distinction between the possible worlds and the impossible ones is a distinction like that between the accidental and essential properties: namely, a distinction that *we* draw rather than one that the world draws. This is the idea I think worth exploring. The thought is that there are no modal facts, but nevertheless we speak truly when we utter sentences in the modal language. How is this? Because there is an underlying basis of facts concerning what abstract worlds there are (and what they represent) which, together with the conventions governing modal language, result in modal statements expressing true propositions. What are the relevant conventions? That a certain subset of the worlds are the possible worlds.

Let us be ersatzist about worlds[3]: every way for the world to be is represented as obtaining by some abstract representation. So we have possible worlds and impossible worlds; but there is no metaphysical difference between the possible worlds and the impossible worlds, it is simply that the conventions governing the term 'possible' pick out a certain subset of the worlds as the possible ones. The possible worlds, on this view, are like the cool people. The cool people are the cool people just because we call them that, not because they instantiate the form of the cool.

[3]I am not going to defend a specific variety of ersatzism here; I hope my position is compatible with whatever account of ersatz worlds you favour.

Likewise, according to me, the possible worlds are the possible worlds simply because we call them the possible worlds. There is no metaphysical difference between them and the impossible worlds; the distinction is a result of how we use our modal language rather than the result of anything in the mind-independent reality.[4]

The difference between this position and the traditional ersatzist position comes down to a difference in the direction of explanation in the biconditional: w is possible iff the situation it says obtains could obtain. For no one should deny that that there *is* the property of being such that you represent a possibility: a property that the possible worlds all have and the impossible worlds all lack. After all, one can simply construct such a property: the property of representing such-and-such a situation as obtaining, or such-and-such a situation as obtaining or . . . etc. My position is simply that this is the property of representing a possibility *because* those worlds are the possible ones, as opposed to the traditional ersatzist thought which is that those worlds are the possible worlds *because* they have this property.

The property of representing a possibility as obtaining is just like coolness. Of course there *is* a property of coolness that all the cool people have and none of the non-cool people have; it is the property of being like such-and-such or like such-and-such or . . . etc. But it is not that the cool people are cool *because* they have this property; rather, this is the property of coolness *because* those people are the cool people.

And of course we are guided by certain considerations in whether we call a certain situation possible (whether it represents something as existing and not existing at the same time, whether it represents something as being red all over and colourless etc) just as we are guided by certain considerations in whether we call a person cool (their good taste in music, their wearing fashionable clothes, their not being a philosopher etc). But we must not think that there is something metaphysically special about these facts. We accord special status to the worlds which represent certain types of situations as obtaining — we call them the possible worlds; but it is *us*, not reality, that puts importance on these considerations; just as it is us, not reality, that puts

[4]The position is very similar to that of Sider [137]. The difference is that for Sider is it a matter of convention what *sentences* are called necessary. It is just a convention, according to Sider, to call certain sentences necessary and not others; so 'necessary' just means something like 'is a truth of logic, or a truth of maths, or an analytic truth, or a truth concerning natural kind membership, or . . . etc' [ibid. p204]. Whereas on the view I am considering the conventions governing modal language determine what *worlds* a proposition has to hold in to be necessary. I don't think a great deal hangs on the difference.

importance on the ways to be cool.

The position here is a quasi-conventionalist position about modality as a whole, not simply the necessary a posteriori or the essential/accidental properties distinction. It is *quasi*-conventionalist because it grants an underlying realm of mind-independent facts; there is no claim being made that the truths we express when we make modal claims are in any sense true in virtue of convention; they, like all truths, are true in virtue of the way things are. But it is quasi-*conventionalist* because there are no brute *modal* facts; we take as primitive non-modal facts and then appeal to convention to account for the true sentences expressed in the modal language. The facts that are taken to be brute are facts concerning the existence of abstracta (the worlds), and then the conventions governing our modal language — namely the conventions governing which of these abstract worlds gets called a possible world — result in our (sometimes) expressing truths when we modalise.

This has potential, I think, to be a very attractive position. We avoid any commitment to brute modal facts. We still have some mind-independent facts holding true brutely; but they are facts concerning the existence of abstracta, which are far more acceptable to be taken as brute than facts concerning what qualitative ways the world could have been like. Facts concerning existence can acceptably be taken as brute if anything can, and do not call out for explanation in the way that modal facts do. But in avoiding primitive modal facts we have not been forced into a non-actualist ontology, like Lewis, nor are we forced into eliminativism. Our folk theory of modality is true; not in virtue of describing some mind-independent modal reality, but in virtue of the existence of abstract representations and the conventions governing our modal language.

Finally, let me say something about how the quasi-conventionalist should respond to Blackburn's dilemma (see chapter 1). Are the facts which make it true that p is necessary, for some necessary proposition p, themselves necessary or contingent according to my theory? Well, there are two facts that are relevant to the necessity of p: the fact that every member of a certain set S of worlds represents p as obtaining and the fact that the members of S are the possible worlds. Is the former fact necessary or contingent? Well I suppose one could go either way, but it is natural to think that it is necessary; what worlds represent to be the case is plausibly essential to them.

But there does not appear to be any sense in which I leave a bad residual 'must'[5] in saying this. For of course the only reason this claim about what the worlds represent is necessary, if it is, is because when it comes to worlds we consider what they represent to be an important feature of them, and hence only count a world as its counterpart if it represents the same thing. So it is not as if there is any unexplained necessity; the necessity of the underlying non-modal facts is given the same quasi-conventionalist treatment any necessary truth is given. What of the facts concerning the conventions? They are contingent of course, for we might have used our words differently; but there is no sense, seemingly, in which this "undermines"[6] the original necessity. We are obviously not forced into saying as a result of this contingency that the necessities could have been false; but nor are we even forced into saying that they could have been contingent either. All that results from the contingency of the facts about convention is that our modal sentences which in fact express modal truths could have expressed falsehoods.

So I am not, in admitting this contingency in the source of modal truth, even committed to denying the S4 axiom. What is wrong with Blackburn's contingency horn, on my account, is the claim that if the source of modal truth involves a contingency then the modal truths could or would have been false had that contingent fact been false. All that follows is that the *sentences* expressing these modal truths could have expressed propositions that are not necessary; but that is because they would be expressing different propositions, not because the necessary propositions they in fact express would have been contingent. The truth or falsity of the propositions those sentences in fact express (i.e. propositions concerning the existence of worlds and their membership in a certain set) is not dependent on what linguistic conventions are in place.

The quasi-conventionalist position promises to be metaphysically attractive in that it has no commitment to brute modal facts or to non-actual ontology, and epistemologically attractive in that it offers an account of why our actual method of discovering what is necessary or possible leads to knowledge. Now of course, more work would have to be done to fully develop this theory. In particular, there is an area in the philosophy of modality that has been badly neglected: namely, an investigation into the function of modal judgements. It may well turn out that this

[5] [19, p635]
[6] ibid.

quasi-conventionalist account of modality will not serve the purposes we want our modal notions to serve. If that is the case, perhaps we should reject the assumption this thesis started with and accept that the modal facts are simply primitive, biting the bullet on the metaphysical and epistemological problems that brings. But I hope I have done enough at least to show that the quasi-conventionalist position merits further investigation.

Bibliography

[1] Roger Ariew and Daniel Garber, editors. *G.W. Leibniz: Philosophical Essays.* Hackitt Publishing Company, Indianapolis and Cambridge, 1989.

[2] D. M. Armstrong. *A Combinatorial Theory of Possibility.* Cambridge University Press, Cambridge, 1989.

[3] David Armstrong. *Nominalism and Realism: Universals and Scientific Realism Volume 1.* Cambridge University Press, Cambridge, 1978.

[4] David Armstrong. *A Theory of Universals: Universals and Scientific Realism Volume 2.* Cambridge University Press, Cambridge, 1978.

[5] David Armstrong. *Universals: An Opinionated Introduction.* Westview Press, Boulder Colorado, 1989.

[6] David Armstrong. *A World of States of Affairs.* Cambridge University Press, Cambridge, 1997.

[7] David Armstrong. *The Mind-Body Problem: An Opinionated Introduction.* Westview Press, Boulder Colorado, 1999.

[8] David Armstrong. Truthmakers for modal truths. In Lillehammer and Rodriguez-Pereyra [78], pages 12–24.

[9] D.M. Armstrong. *Truth and Truthmakers.* Cambridge University Press, Cambridge, 2004.

[10] D.F. Austin, editor. *Philosophical Analysis.* Kluwer, Dordrecht, 1988.

[11] J.L. Austin. Truth. In Pitcher [103], pages 18–31.

[12] John Bacon, Keith Campbell, and Lloyd Reinhardt, editors. *Ontology, Causality and Mind.* Cambridge University Press, Cambridge, 1993.

[13] Tom Baldwin. There might be nothing. *Analysis*, 56:231–38, 1996.

[14] Paul Benacerraf. Mathematical truth. *Journal of Philosophy*, 70:661–679, 1973.

[15] John Bigelow. *The reality of numbers: a physicalist's philosophy of mathematics.* Oxford University Press, Oxford, 1988.

[16] Max Black. The identity of indiscernibles. In *Problems of Analysis* [17], pages 80–92.

[17] Max Black. *Problems of Analysis.* Routledge and Kegan Paul, London, 1954.

[18] Max Black. Possibility. *Journal of Philosophy*, 57:117–126, 1960.

[19] Simon Blackburn. Morals and modals. In Kim and Sosa [58], pages 634–648.

[20] George Boolos. *Logic, Logic, and Logic.* Harvard University Press, Cambridge, MA, 1998.

[21] George Boolos. To be is to be a value of a variable (or to be some values of some variables). [20], pages 54–72.

[22] Phillip Bricker. Island universes and the analysis of modality. In Preyer and Siebelt [108], pages 27–55.

[23] Richard Cartwright. Some remarks on essentialism. *Journal of Philosophy*, 65:615–626, 1968.

[24] Richard Cartwright. *Philosophical Essays.* MIT Press, Cambridge, 1987.

[25] Richard Cartwright. Scattered objects. [24], pages 171–186.

[26] Charles Chihara. *The Worlds of Possibility.* Oxford University Press, Oxford, 1998.

[27] Geraldine Coggins. World and object: Metaphysical nihilism and three accounts of worlds. *Proceedings of the Aristotelian Society*, CIII(3):353–360, 2003.

[28] Tim Crane and Katalin Farkas, editors. *Metaphysics: a guide and anthology.* Oxford University Press, Oxford, 2004.

[29] Donald Davidson and Jaakko Hintikka, editors. *Words and Objections: Essays on the work of W. V. Quine.* D. Riedel Publishing Company, Dordrecht, Holland, 1969.

[30] John Divers. The analysis of possibility and the possibility of analysis. *Proceedings of the Aristotelian society*, 97:141–160, 1997.

[31] John Divers. A genuine realist theory of advanced modalizing. *Mind*, 108:217–239, 1999.

[32] John Divers. *Possible Worlds.* Routledge, London, 2002.

[33] John Divers. Agnosticism about other worlds: A new anti-realist programme in modality. *Philosophy and Phenomenological Research*, 2004.

[34] John Divers. On the state of modalizing de re: Quine after Lewis. forthcoming.

[35] John Divers and Joseph Melia. The analytic limit of genuine modal realism. *Mind*, 111(441):15–36, 2002.

[36] Michael Dummett. Wittgenstein's philosophy of mathematics. *Philosophical Review*, 68:324–348, 1959.

[37] David Efird and Tom Stoneham. The subtraction argument for metaphysical nihilism. *The Journal of Philosophy*, forthcoming, 2005.

[38] Kit Fine. Essence and modality. *Philosophical Perspectives*, 8:1–16, 1994.

[39] Kit Fine. The varieties of necessity. In *Conceivability and Possibility* [143], pages 253–281.

[40] Graeme Forbes. *The Metaphysics of Modality.* Oxford Clarendon Press, Oxford, 1985.

[41] Graeme Forbes. In defense of absolute essentialism. *Midwest Studies in Philosophy*, 11:3–31, 1986.

[42] Peter Forrest. Ways worlds could be. *Australasian Journal of Philosophy*, 64(1):15–24, 1986.

[43] John Fox. Truthmaker. *Australasian Journal of Philosophy*, 65:188–207, 1987.

[44] Andre Gallois. *Occasions of Identity.* Oxford University Press, Oxford, 1998.

[45] Allan Gibbard. Contingent identity. *Journal of Philosophical Logic*, 4:187–221, 1975.

[46] Bob Hale. *Abstract objects.* Oxford University Press, Oxford, 1987.

[47] Bob Hale. The source of necessity. *Philosophical Perspectives*, 16:299–319, 2002.

[48] Bob Hale. Knowledge of possibility and necessity. *Proceedings of the Aristotelian Society*, CIII(1):1–20, 2003.

[49] Bob Hale. The necessity of identity. *Logica Handbook*, pages 21–33, 2003.

[50] Bob Hale. Putnam's retreat: Some reflections on Hilary Putnam's changing views about metaphysical necessity. *Midwest Studies in Philosophy*, 28:351–378, 2004.

[51] John Hawthorne and Tamr Szabo Gendler. Origin essentilaism: The arguments reconsidered. *Mind*, 109:285–298, 2000.

[52] A. Hazen. Counterpart-theoretic semantics for modal logic. *Journal of Philosophy*, 76:319–338, 1979.

[53] John Heil. Truth making and entailment. *Logique et Analyse*, 43(169-170), 2000.

[54] John Heil. *From an Ontological Point of View.* Oxford University Press, Oxford, 2003.

[55] G.E. Hughes and M.J. Cresswell. *A New Introduction to Modal Logic.* Routledge, London and New York, 1996.

[56] Frank Jackson and Graham Priest, editors. *Lewisian Themes.* Clarendon Press, Oxford, 2004.

[57] Michael Jubien. Problems with possible worlds. In Austin [10], pages 299–322.

[58] Jaegwon Kim and Ernest Sosa, editors. *Metaphysics: An Anthology.* Blackwell, Oxford, 1999.

[59] Saul Kripke. Identity and necessity. In Moore [94], pages 162–191.

[60] Saul A. Kripke. *Naming and necessity.* Harvard University Press, Cambridge, MA, 1980.

[61] David Vander Laan. Counterpossibles and similarity. In Jackson and Priest [56], pages 258–275.

[62] David Lewis. Counterpart theory and quantified modal logic. In *Philosophical Papers Volume 1* [64], pages 26–39.

[63] David Lewis. Counterparts of persons and their bodies. In *Philosophical Papers Volume 1* [64], pages 47–54.

[64] David Lewis. *Philosophical Papers Volume 1.* Oxford University Press, Oxford, 1983.

[65] David Lewis. Postscripts to 'Counterpart theory and quantified modal logic'. In *Philosophical Papers Volume 1* [64], pages 39–46.

[66] David Lewis. *On the Plurality of Worlds.* Blackwell, Oxford, 1986.

[67] David Lewis. *Parts of Classes.* Blackwell, Oxford, 1992.

[68] David Lewis. Against structural universals. In *Papers in Metaphysics and Epistemology* [71], pages 78–107.

[69] David Lewis. A comment on Armstrong and Forrest. In *Papers in Metaphysics and Epistemology* [71], pages 108–110.

[70] David Lewis. Noneism or allism? In *Papers in Metaphysics and Epistemology* [71], pages 152–163.

[71] David Lewis. *Papers in Metaphysics and Epistemology.* Cambridge University Press, Cambridge, 1999.

[72] David Lewis. A world of truthmakers? In *Papers in Metaphysics and Epistemology* [71], pages 215–220.

[73] David Lewis. Forget about the 'correspondence theory of truth'. *Analysis*, 61(4):275–280, 2001.

[74] David Lewis. Truthmaking and difference-making. *Nous*, 35(4):602–615, 2001.

[75] David Lewis. Things qua truthmakers. In Lillehammer and Rodriguez-Pereyra [78], pages 25–38.

[76] David Lewis and Rae Langton. Defining 'intrinsic'. In *Papers in Metaphysics and Epistemology* [71], pages 116–132.

[77] David Lewis and Gideon Rosen. Postscript to 'Things qua truthmakers': negative existentials. In Lillehammer and Rodriguez-Pereyra [78], pages 39–42.

[78] Hallvard Lillehammer and Gonzalo Rodriguez-Pereyra, editors. *Real Metaphysics*. Routledge, London, 2003.

[79] Leonard Linsky, editor. *Reference and Modality*. Oxford University Press, London, 1971.

[80] Michael J. Loux and Dean W. Zimmerman, editors. *The Oxford Handbook of Metaphysics*. Oxford University Press, Oxford, 2003.

[81] E.J. Lowe. On the alleged necessity of true identity statements. *Mind*, 91(364):579–584, 1982.

[82] E.J. Lowe. Why is there anything at all? *Proceedings of the Aristotelian Society Supplementary Volume*, 70:111–120, 1996.

[83] E.J. Lowe. *The Possibility of Metaphysics*. Clarendon Press, Oxford, 1998.

[84] E.J. Lowe. Metaphysical nihilism and the subtraction argument. *Analysis*, 62(1):62–73, 2002.

[85] William G. Lycan. Pot bites kettle: A reply to miller. *Australasian Journal of Philosophy*, 69:212–213, 1991.

[86] William G. Lycan. Two - no, three - concepts of possible worlds. *Proceedings of the Aristotelian Society*, 91:215–227, 1991.

[87] Kris McDaniel. Modal realism with overlap. In Jackson and Priest [56], pages 140–155.

[88] Joseph Melia. A note on lewis' ontology. *Analysis*, 52:191–192, 1992.

[89] D.H. Mellor. *Real Time II.* Routledge, London and New York, 1998.

[90] D.H. Mellor. Real metaphysics: Replies. In Lillehammer and Rodriguez-Pereyra [78], pages 212–238.

[91] Richard B. Miller. Dog bites man: A defence of modal realism. *Australasian Journal of Philosophy*, 67:476–478, 1989.

[92] Richard B. Miller. There is nothing magical about possible worlds. *Mind*, 99:453–457, 1990.

[93] Richard B. Miller. Genuine modal realism: Still the only non-circular game in town. *Australasian Journal of Philosophy*, 71:159–160, 1993.

[94] A.W. Moore, editor. *Meaning and Reference.* Oxford University Press, Oxford, 1993.

[95] G.E. Moore. *Principia Ethica.* Cambridge University Press, Cambridge, 1903.

[96] Daniel Nolan. Impossible worlds: A modest approach. *Notre Dame Journal of Formal Logic*, 38(4):535–572, 1997.

[97] Josh Parsons. There is no truthmaker argument against nominalism. *Australasian Journal of Philosophy*, 77(3):325–334, 1999.

[98] Josh Parsons. Must a four-dimensionalist believe in temporal parts? *Monist*, 83(3):399–418, 2000.

[99] Alexander Paseau. Why the subtraction argument does not add up. *Analysis*, 62(1):73–75, 2002.

[100] L.A. Paul. The context of essence. In Jackson and Priest [56], pages 181–195.

[101] Christopher Peacocke. Metaphysical necessity: Understanding, truth and epistemology. *Mind*, 106:521–574, 1997.

[102] Christopher Peacocke. *Being Known.* Clarendon Press, Oxford, 1999.

[103] George Pitcher, editor. *Truth.* Prentice-Hall, Englewood Cliffs, N.J., 1964.

[104] Alvin Plantinga. *The nature of necessity.* Clarendon Press, Oxford, 1974.

[105] Alvin Plantinga. *Essays in the Metaphysics of Modality.* Oxford University Press, Oxford, 2003.

[106] Alvin Plantinga. Two concepts of modality: Modal realism and modal reductionism. [105], pages 192–228.

[107] Robin Le Poidevin and Murray MacBeath. *The Philosophy of Time.* OxfordUniversity-Press, Oxford, 1993.

[108] G Preyer and F Siebelt, editors. *Reality and Humean Supervenience: Essays on the Philosophy of David Lewis.* Rowman and Littlefield, Lanham, Maryland, 2001.

[109] Hilary Putnam. The meaning of 'meaning'. In *Mind, language, and reality — philosophical papers, vol.2* [110], pages 215–271.

[110] Hilary Putnam. *Mind, language, and reality — philosophical papers, vol.2.* Cambridge University Press, Cambridge, 1975.

[111] Hilary Putnam. Possibility and necessity. In *Realism and reason: collected philosophical papers, volume III* [112].

[112] Hilary Putnam. *Realism and reason: collected philosophical papers, volume III.* Cambridge University Press, Cambridge, 1983.

[113] Hilary Putnam. Is water necessarily H2O? In *Realism with a human face* [114].

[114] Hilary Putnam. *Realism with a human face.* Harvard University Press, Cambridge, MA, 1990.

[115] W. V. Quine. Two dogmas of empiricism. *Philosophical Review*, 60:20–43, 1951.

[116] W. V. Quine. *From a logical point of view.* Harvard University Press, Cambridge, MA, 1953.

[117] W. V. Quine. Reference and modality. In *From a logical point of view* [116], pages 139–159.

[118] W. V. Quine. *Word and Object.* MIT Press, Cambridge, MA, 1960.

[119] W. V. Quine. *Ontological relativity and other essays.* Columbia University Press, New York, 1969.

[120] W. V. Quine. *The ways of paradox and other essays.* Harvard University Press, Cambridge, MA, 1976.

[121] W.V. Quine. Three grades of modal involvment. In *The ways of paradox and other essays* [120], pages 158–176.

[122] Stephen Read. Truthmakers and the disjunction thesis. *Mind*, 109(433):67–80, 2000.

[123] Greg Restall. Truthmakers, entailment and necessity. *Australasian Journal of Philosophy*, 74:331–340, 1996.

[124] Tom Richards. The worlds of David Lewis. *Australasian Journal of Philosophy*, 53:105–118, 1975.

[125] Teresa Robertson. Possibilities and the argumetns for origin essentialism. *Mind*, 107:729–749, 1998.

[126] Teresa Robertson. Essentialism: Origin and order. *Mind*, 108:299–307, 2000.

[127] Gonzalo Rodriguez-Pereyra. There might be nothing: the subtraction argument improved. *Analysis*, 57(3), 1997.

[128] Gonzalo Rodriguez-Pereyra. Metaphysical nihilism defended: reply to Lowe and Paseau. *Analysis*, 62(2):172–180, 2002.

[129] Gonzalo Rodriguez-Pereyra. Modal realism and metaphysical nihilism. *Mind*, 113:1–22, 2004.

[130] Guy Rohbraugh and Louis DeRossett. A new route to the necessity of origin. *Mind*, 113:705–725, 2004.

[131] Nathan Salmon. How not to derive essentialism from the theory of reference. *Journal of Philosophy*, 76:703–725, 1979.

[132] Nathan Salmon. *Reference and Essence.* Princeton University Press, Princeton, 1981.

[133] Scott A. Shalkowski. The ontological ground of the alethic modality. *Philosophical Review*, 103:669–688, 1994.

[134] Theodore Sider. All the world's a stage. *Australasian Journal of Philosophy*, 74:433–453, 1996.

[135] Theodore Sider. *Four-dimensionalism.* Oxford University Press, Oxford, 2001.

[136] Theodore Sider. Maximality and intrinsic properties. *Philosophy and Phenomenological Research*, 63(2):357–364, 2001.

[137] Theodore Sider. Reductive theories of modality. In Loux and Zimmerman [80], pages 180–208.

[138] J.J.C. Smart. Sensations and brain processes. *Philosophical Review*, (68), 1959.

[139] Arthur F Smullyan. Modality and description. In Linsky [79], pages 35–43.

[140] Ernest Sosa and Jaegwon Kim. *Epistemology: An Anthology.* Blackwell, Oxford, 2000.

[141] Robert Stalnaker. Possible worlds. [142], pages 25–39.

[142] Robert Stalnaker. *Ways a World Might Be.* Clarendon Press, Oxford, 2003.

[143] John Hawthorne Tamar Szabo Gendler. *Conceivability and Possibility.* Clarendon Press, Oxford, 2002.

[144] Judith Jarvis Thomson. The statue and the clay. *Noûs*, 32:149–173, 1998.

[145] James E Tomberlin and Peter van Inwagen, editors. *Alvin Plantinga.* Kluwer, The Netherlands, 1985.

[146] Peter van Inwagen. Plantinga on trans-world identity. In Tomberlin and van Inwagen [145], pages 101–120.

[147] Peter van Inwagen. Two concepts of possible worlds. *Midwest Studies in Philosophy*, 11:185–213, 1986.

[148] Peter van Inwagen and Dean Zimmerman, editors. *Metaphysics: The Big Questions*. Blackwell, Oxford, 1998.

[149] Timothy Williamson. Existence and contingency. *Proceedings of the Aristotelian Society Supplementary Volume*, 73:181–203, 1999.

[150] Stephen Yablo. Is conceivability a guide to possibility? *Philosophy and Phenomenological Research*, LIII(1):1–42, 1993.

[151] Stephen Yablo. Intrinsicness. *Philosophical Topics*, 26:479–505, 1999.